NASA CR- 134708

Final Report

APPLICATION OF SYSTEM THEORY TO POWER PROCESSING PROBLEMS

Roger W. Brockett and Jonathan R. Wood

HARVARD UNIVERSITY
Division of Engineering and Applied Physics
Cambridge, Massachusetts 02138

prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA Lewis Research Center

Contract NGR 22-007-172

1. Report No. NASA CR-134708	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle FINAL REPORT APPLICATION OF SYSTEM THEORY TO POWER PROCESSING PROBLEMS		5. Report Date 6. Performing Organization Code
7. Author(s) Roger W. Brockett and Jonathan R. Wood		8. Performing Organization Report No.
9. Performing Organization Name and Address Harvard University Division of Engineering and Applied Physics Cambridge, Massachusetts 02138		10. Work Unit No. 120-60-03(YOS-4224) 11. Contract or Grant No. NGR 22-007-172
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		13. Type of Report and Period Covered FINAL 1970-1974 14. Sponsoring Agency Code

15. Supplementary Notes

Project Manager, Vincent R. Lalli, Spacecraft Technology Division, NASA Lewis Research Center, Cleveland, Ohio

16. Abstract

Rewrite

This report summarizes our efforts to apply system theory to power processing problems. In Chapter 1 we show that input-output models of the form

$$\dot{X}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)x(t); \quad Y(t) = Cx(t)$$

discussed

Are are capable of representing a wide variety of highly nonlinear problems of practical importance and, at the same time, represent a class of systems for which a fairly detailed structural analysis can be made. ~~In Chapter 2 we discuss~~ the problem areas in differential equations with methods and tools from control theory which have been useful in obtaining new results in differential equation theory. ~~In the treatment of such topics as the existence of geodesics and obtaining expressions for the spectrum of the Laplacian, one finds that the Lie algebras and groups are closely connected.~~ I~~n Chapter 3 we discuss~~ some practical methods of analysis for time invariant electrical networks which contain controlled switches. exp. In~~ Chapter 4 we discuss~~ the overall plan that this research followed comparing the desired objectives with those achieved and define a new area of work to improve control system reliability through design simplifications.

are presented along with

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Power Conditioning Electronics Frequency Domain Stability Criteria Lie Algebras Reliability	Unclassified - unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	101	

* For sale by the National Technical Information Service, Springfield, Virginia 22151

NASA-C-168 (Rev. 6-71)

NASA CR-134708

FINAL REPORT

APPLICATION OF SYSTEM THEORY TO POWER PROCESSING PROBLEMS

by Roger W. Brockett and Jonathan R. Wood

HARVARD UNIVERSITY

Division of Engineering and Applied Physics

Cambridge, Massachusetts 02138

prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA Lewis Research Center

Grant NGR 22-007-172

Page intentionally left blank

CONTENTS

NOTATION

$x \in A$ means that x is a member of the set A.

$A \Rightarrow B$ means that A implies B.

$\{x | x \text{ has property } A\}$ denotes the set of all x such that x has property A.

$G: A \to B$ means that the operator (or function) G maps the set A into the set B.

$\dot{x}$ denotes $\frac{dx}{dt}$.

$\mathrm{Re}\, z$ denotes the real part of the complex number z.

$\mathrm{Im}\, z$ denotes the imaginary part of z.

$x < \infty$ means that x is finite.

$x(t) \equiv a$ means that $x(t) = a$ for all t.

$\lim_{t \to \infty} x(t) = a$, or $x(t) \to a$ as $t \to \infty$, means that for all $\eta > 0$ there is a T such that $|x(t)-a| < \eta$ for all $t \geq T$.

$\sup_{x \in A} x$ denotes the supremum (or least upper bound) of the set of numbers A, i.e. the least number y such that $x \leq y$ for all $x \in A$.

$\inf_{x \in A} x$ denotes the infimum (or greatest lower bound) of the set of numbers A, i.e. the greatest number z such that $x \geq z$ for all $x \in A$.

$\underset{(-a,b)}{\mathrm{stp}}\ \sigma$ is the function on the real line defined by

$$\underset{(-a,b)}{\mathrm{stp}}\ \sigma = \begin{cases} -a, & \sigma < 0 \\ 0, & \sigma = 0 \\ b, & \sigma > 0 \end{cases}.$$

$\underset{(-a,b)}{\mathrm{Sod}}\ \sigma$ is the function defined by

$$\underset{(-a,b)}{\mathrm{Sod}}\ \sigma = \sigma \underset{(-a,b)}{\mathrm{stp}}\ \sigma .$$

$\mathbb{R}$ denotes the set of all real numbers.

$\mathbb{R}^{m \times n}$ denotes the set of all mxn real matrices.

$\mathbb{R}^n$ denotes the set of all n-dimensional real vectors.

A matrix is denoted $\underline{A}$, a column- or row-vector is denoted $\underline{b}$.

ON THE ALGEBRAIC STRUCTURE OF BILINEAR SYSTEMS

Roger W. Brockett

Harvard University
Cambridge, Massachusetts

Abstract

DATA MISSING

Preliminary Ideas

Following a general theme in the mathematical theory of model building, our concern here is with the relationship between external (often emperical) descriptions of dynamic system, and internal (for us a description in terms of differential equations) descriptions of the model. We refer to the latter as a realization of an input-output system. The system itself is thought of as a collection of input-output pairs.

*This work was supported in part by the U.S. Office of Naval Research under the Joint Services Electronics Program by Contract N00014-67-A-0298-0006 and by the National Aeronautics and Space Administration under Grant NGR 22-007-172.

We want to describe a theory which is general enough to treat systems of the form

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t)+\sum_{i=1}^{m} u_i(t)b_i \quad ; \quad y(t) = c[x(t)]$$

where A and B_i are square matrices, the b_i are column vectors and c[x] is a finite power series. This departure from linear systems, i.e. systems for which the $u_i(t)B_i$ terms are absent, and c is linear, is justified on the grounds that a number of practical control problems can only be modelled successfully if the multiplicative control and output nonlinearity are present. The reasons for this

DATA MISSING

For bilinear models the tools of linear algebra are no longer enough. There is a simple explanation of this fact. In order to decompose the system equations as completely as possible, it is necessary to develop canonical forms for a set of matrices which admit both linear operations and a type of multiplication. A form which is convenient relative to the vector space structure of the set of matrices typically is not well behaved relative to the multiplicative structure and conversely. To sort this all out requires more than just linear algebra. For reasons having to do with controllability, the useful multiplication rule is [A,B] = AB-BA. The study of bilinear systems is intimately connected, therefore with the study of sets of matrices which are closed under vector space operation and also the above multiplication. These objects form Lie algebras and if we are to make reasonable progress in understanding bilinear systems, this theory cannot be avoided.

Examples leading to bilinear constraints include those where energy is to be conserved. If x must satisfy

$$x'Qx = 1$$

then we may model a controlled system by

$$\dot{x}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)x(t)$$

where $QA + A'Q = 0$ and $QB_i + B_i'Q = 0$.

Higher order constraints can also be accommodated. Let $V_1, V_2, \ldots, V_r$ and W be vector spaces over the same field. A map

$$\phi: V_1 \times V_2 \times \ldots \times V_r \to W$$

is called multilinear if it satisfies, for all α and β in the field and all $i = 1,2,\ldots r$.

$$\phi(v_1,v_2,\ldots,\alpha v_i+\beta v_i',\ldots,v_{r-1},v_r)$$
$$= \alpha\phi(v_1,v_2,\ldots,v_i,\ldots,v_{r-1},v_r)+\beta\phi(v_1,v_2,\ldots,v_i',\ldots,v_{r-1},v_r)$$

Given a multilinear form $\phi : \mathscr{R}^n \times \mathscr{R}^n \times \ldots \times \mathscr{R}^n \to \mathscr{R}$, suppose the constraint to be satisfied by x is

$$\phi(x,x,\ldots,x) = 1$$

Let the equations of motion be

$$\dot{x}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)x(t)$$

This imposes the conditions on A and B_i

$$L(Ax,x,\ldots x)+L(x,Ax,\ldots x)+\ldots+L(x,x,\ldots Ax) = 0$$
$$L(B_ix,x,\ldots x) + L(x,B_ix,\ldots x)+\ldots+L(x,x,\ldots B_ix) = 0$$

Specific instances which require both the additive and the multiplicative terms have been given in the literature [1]. One large class of problems of this type arise in the study of switched electrical networks, examples of which appear in [2] and [3]. The bilinear form is of basic importance in certain problems having a geometrical component due to the Frenet-Serret formulas for curves in a 3-dimensional space.

The Basic Bilinear Model

We want to show that a large class of input-output models can be reduced to the form

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t) \; ; \; y(t) = Cx(t) \qquad \text{(I)}$$

where x is an n-tuple, y is a q-tuple and A, $\{B_i\}$ and C are matrices of appropriate dimensions.

We begin with a simple observation. (Compare with [2] section 7 and [3] section 4.)

Theorem 1: Any input-output map which can be realized by a set of equations of the form

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t)+\sum_{i=1}^{m} u_i(t)b_i \quad ; \quad y(t) = Cx(t)$$

can be realized by a set of equations of the form

$$\dot{z}(t) = (F+\sum_{i=1}^{m} u_i(t)G_i)z(t) \quad ; \quad y(t) = Hz(t) \quad \text{(I')}$$

Proof: Let F and G_i be defined by adding a single extra row and column to A, and B_i respectively

$$F = \begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix} \qquad G_i = \begin{bmatrix} 0 & 0 \\ b_i & B_i \end{bmatrix}$$

Let z and H be given by

$$z = \begin{bmatrix} 1 \\ x \end{bmatrix} \qquad H = [0, C]$$

It is immediate that the z-system defines the same input-output map as the x-system. □

The second result is a little more involved. It shows that nonlinear output maps can be reduced to linear forms provided they are of the finite power series type. This is the kind of result that has no counter part in linear theory and points out the great flexibility inherit in the bilinear model. The basis for the result is the observation (which goes all the way back to the thesis of A.M. Liapunov) is that if x satisfies a linear equation then x(t)x'(t) satisfies one also. Thus in our case if x satisfies

(I) then (prime denotes transpose)

$$\frac{d}{dt}x(t)x'(t)=(A+\sum_{i=1}^{m}u_i(t)B_i)x(t)x'(t)+x(t)x'(t)(A+\sum_{i=1}^{m}u_i(t)B)'$$

which is an equation of the bilinear type! That is, there exist matrices $A^{[2]}$ and $B_i^{[2]}$ such that $z=(x_1^2,x_1x_2,x_1x_3,\ldots x_2^2,x_2x_3,\ldots,x_n^2)$ satisfies

$$\dot{z}(t) = (A^{[2]} + \sum_{i=1}^{m} u_i(t)B_i^{[2]})z(t)$$

Of course $A^{[2]}$ and $B_i^{[2]}$ are derived from A and B_i, respectively. One can be more explicit using Kronecker product relationships and the theory of symmetric tensors [4]. The same is true not only for $\{x_ix_j\}$ but also $\{x_ix_jx_k\}$ etc. as is easily verified. Thus associated with each bilinear equation is a countable collection of bilinear systems. The mth entry in this collection being the bilinear equation for the mth-degree forms in x. It can be taken to be of dimension equal to the number of linearly independent m-forms in n variables, i.e. $n(n+1)\ldots(n+m-1)/2.3\cdot\ldots m$. We indicate the vector consisting of these forms (ordered lexographically, for the sake of definiteness) by $x^{[m]}$.

Theorem 2: Any input-output map which can be realized in the form

$$\dot{x}(t)=(A+\sum_{i=1}^{m}u_i(t)B_i)x(t) \quad ; \quad y(t)=\sum_{p=1}^{q}L_p(x(t),x(t),\ldots,x(t))$$

where L_p is a p-linear map can be realized in the form

$$\dot{z}(t) = (F+\sum_{i=1}^{m}u_i(t)G_i)z(t) \quad ; \quad y(t) = Hz(t)$$

Proof: It is clear from the previous remarks that if x satisfies a bilinear state equation then so does $x^{[m]}$. Thus we can write an equation of the form

$$\dot{z}(t) = [\tilde{A} + \sum_{i=1}^{m} u_i(t)\tilde{B}_i]z(t)$$

where z is defined by

$$z' = (x,x^{[2]},\ldots x^{[q]})$$

and $[\tilde{A} + \sum_{i=1}^{m} u_i(t)\tilde{B}_i]$ is given by

$$\begin{bmatrix} A+u_i(t)B_i & 0 & \cdots & 0 \\ 0 & A^{[2]}+u_i(t)B_i^{[2]} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & & A^{[q]}+u_i(t)B_i^{[q]} \end{bmatrix}$$

Now y is a linear combination of the components of z since it is multilinear in the components of x. □

Example: The reader may verify that the input-output system defined by

$$\ddot{x} = u \quad ; \quad y = x^2$$

is represented by

$$\frac{d}{dt}\begin{bmatrix} 1 \\ x \\ \dot{x} \\ x^2 \\ \dot{x}x \\ \dot{x}^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ u & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & u & 0 & 0 & 0 & 1 \\ 0 & 0 & 2u & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ x \\ \dot{x} \\ x^2 \\ \dot{x}x \\ \dot{x}^2 \end{bmatrix}$$

$$y = [0 \;\; 0 \;\; 0 \;\; 1 \;\; 0 \;\; 0]x$$

System Interconnection

We say that two bilinear systems (I) and (I') are interconnected in parallel to get the single system if we simply add these outputs. That is, the equations for the parallel inter-connection are

$$\dot{x}(t)=(A+\sum_{i=1}^{n} u_i(t)B_i)x(t)$$

$$; \; y(t) = Cx(t)+Hz(t)$$

$$\dot{z}(t)=(F+\sum_{i=1}^{n} u_i(t)G_i)z(t)$$

Clearly this is defined only if the dimensionality of the

input spaces of I and I' are the same and the dimensionality of the output spaces of the two systems are the same.

We say that two bilinear systems are interconnected in series with (I') following (I) if the input to (I') is equated to the output of (I) the equations for the series interconnection are

$$\dot{x}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)x(t)$$
$$; \; y(t) = Hz(t)$$
$$\dot{z}(t) = (F + \sum_{i=1}^{m}(Cx)_i(t)G_i)z(t)$$

Clearly a series connection is possible if the dimension of the output of the first system equals the dimension of the input of the second.

Remark: If the series interconnection of two input-output systems having bilinear realizations is defined then the system which results from parallel interconnection has a bilinear realization. If the series connection of a system having a bilinear realization followed by a system having a linear realization is defined, then the system which results from series interconnection has a bilinear realization.

We have not been able to determine if the class of bilinear realizations is closed under series interconnection.

The Canonical Form

The existance of the Jordan normal form for a linear map of $\mathscr{R}^n$ into $\mathscr{R}^n$ gives rise to the "diagonal" or "partial fraction" realization for linear systems. This is important because in certain senses the Jordan form displays the maximum degree decoupling which is possible. We want to describe the analogous situation for bilinear systems. As might be expected, the results cannot be based on the tools of linear algebra alone.

In view of the results of section 2, we are content to consider hence forth systems which have realizations in the form of equation (I).

We call two realizations

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i x(t) \quad ; \quad y(t) = Cx(t) \tag{I}$$

and

$$\dot{z}(t) = (F+\sum_{i=1}^{m} u_i(t)G)x(t) \quad ; \quad y(t) = Hz(t) \tag{I'}$$

equivalent if there exists a nonsingular P such that $PAP^{-1} = F$ and $PB_iP^{-1} = G_i$ and $CP^{-1} = H$.

We call a realization in the form (I) irreducible if there is no nonsingular P such that

$$PAP^{-1} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} \quad PB_iP^{-1} = \begin{bmatrix} \tilde{B}^i_{11} & 0 \\ \tilde{B}^i_{21} & \tilde{B}_{22} \end{bmatrix}$$

where $\tilde{A}_{11}$ and $\tilde{B}^i_{11}$ are square matrices, all of the same dimension. That is, for no choice of basis is the realization in block triangular form. Otherwise we call it reducible. A reducible realization is said to be completely reducible if it can be put in block diagonal form (as opposed to block triangular form) with each block being irreducible. A realization of the form (I) said to be equivalent to a triangular realization if there exists a nonsingular P (possibly complex) such that PAP^{-1} and PB_iP^{-1} are lower triangular. (Including the possibility of nonzero elements on the diagonal.) We call it strictly triangular if there exists P such that PAP^{-1} and PB_iP^{-1} are strictly lower triangular. (No nonzero elements on the diagonal.)

If a system is reducible then there are nontrivial invariant subspaces for the collection of matrices $\{A,B_i\}$. Let V_1 be one of these which is of smallest (positive) dimension. (There may be many, pick any one.) Let V_2 be a smallest invariant subspace properly containing V_1. Let V_3 be a smallest invariant subspace properly containing V_2, etc. Let $n_i = \dim V_i$. Pick a basis such that the first n_1 elements span the space V_1, the first n_2 elements span V_2, etc. Relative to this basis the matrices A and B_i take the block triangular form

$$A = \begin{bmatrix} A_{11} & 0 & 0 & \cdots \\ A_{12} & A_{22} & 0 & \cdots \\ A_{13} & A_{23} & A_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} \quad B_i = \begin{bmatrix} B^i_{11} & 0 & 0 & \cdots \\ B^i_{12} & B^i_{22} & 0 & \cdots \\ B^i_{13} & B^i_{23} & B^i_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

Each of the collection of block diagonals $\{A_{kk}, B^i_{kk}\}$ are irreducible and the Jordan-Holder Theorem insures that these representations are unique in that regardless of how the invariant subspaces are chosen, the construction will lead to an equivalent collection of irreducible diagonal blocks. (They may occur in a different order depending on the choice of subspace, of course.) We collect these observations in a theorem. (See, e.g. Samelson [4] page 12 for a sketch of a proof.)

Theorem 3: Every bilinear realization (I) is equivalent to one in which the A and B_i matrices are in block triangular form with the diagonal blocks being irreducible. Moreover if (A,B_i,C) and (F,G_i,H) are two equivalent realizations in block triangular form with irreducible blocks on the diagonal then there is a permutation π and nonsingular matrices P_k such that the diagonal blocks are related by

$$P_k A_{kk} P_k^{-1} = F_{\pi(k)\pi(k)} \quad ; \quad P_k B_{kk} P_k^{-1} = G_{\pi(k)\pi(k)}$$

We will say that an input-output system displayed according to the above recipe is in a reduced form.

Controllability

A detailed study of the controllability properties of bilinear and even more general systems, has been made in the recent literature. References [5] - [8] contain many interesting results. For our present purposes section 7 of [2] and section 6 of [8] are relevant.

In reference [2] it is shown that if A is zero, or if a certain commutation condition is satisfied, then the reachable set for

$$\dot{x}(t) = (A + \sum_{i=1}^{m} u_i(t) B_i) x(t) \; ; \; y(t) = Cx(t) \; ; \; x(0) = x_0 \quad \text{(I)}$$

is easily computed. However Jurdjevic and Sussmann [8] have shown that the reachable set for (I) contains an open subset of the set reachable for

$$\dot{x}(t) = (v(t)A + \sum_{i=1}^{m} u_i(t)B_i)x(t) \ ; \ y(t) = Cx(t) \ ; \ x(0) = x_o \quad \text{(II)}$$

From this fact it is easy to show that the reachable set for (I) is confined to a subspace if and only if the reachable set for (II) is confined to the same subspace. We omit the details but make explicit use of this result below.

Theorem 4: The reachable set for (I) is confined to a subspace if and only if there exists a nonsingular P such that

$$PAP^{-1} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}$$

$$PB_iP^{-1} = \begin{bmatrix} \tilde{B}^i_{11} & 0 \\ \tilde{B}^i_{21} & \tilde{B}^i_{22} \end{bmatrix}$$

$$P^{-1}x_o = \begin{bmatrix} 0 \\ \tilde{x}_o \end{bmatrix}$$

where the 0 blocks are all of the same dimension.

Proof: If there exists such a P then clearly the reachable set is confined to the subspace consisting of those vectors whose upper portion is zero.

Suppose the reachable set of (I) is confined to a subspace. Then by our remarks above the reachable set for (II) is confined to a subspace. But from the results of section 7 of [2] we see that this implies that A and B_i can be simultaneously block triangularized. □

Remark: Notice that the set of matrices $\{A,B_i\}$ can be simultaneously triangularized if and only if one can simultaneously triangularize the larger set obtained from $\{A,B_i\}$ by adjoining all linear combination products of any two elements, products of products, etc. More precisely, we define $\{A,B_i\}_{AA}$ to be the smallest vector space of matrices which contains $\{A,B_i\}$ and is closed under multiplication by elements of $\{A,B_i\}$. This larger set is called the associative algebra.

generated by $\{A,B_i\}$. The condition of Theorem 4 can be stated as requiring that x_0 should not belong to any subset of $\mathscr{R}^n$ which is invariant with respect to multiplication by elements of the associative algebra. This statement is close to the familiar $(B,AB,A^2B,\ldots)$ test for controllability.

Theorem 5: Any input-output map which can be realized by a bilinear system can be realized by one for which the reachable set is not confined to a linear subspace.

Proof: Use Theorem 4. If the reachable set is confined to a subspace find the P which effects the decomposition for Theorem 4. Delete the top block, then the input-output map is the same but the state is not confined to a linear subspace. □

Observability

We will say that two starting states, x_o and x_1 of the system

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t) \quad ; \quad y(t) = Cx(t) \qquad \text{(I)}$$

are indistinguishable if for all inputs u, the response y is the same. This follows our approach in [2] where more general output maps are considered. We start off the analog of Theorem 4.

Theorem 6: The system (I) has no indistinguishable states if and only if there exists no nonsingular P such that CP^{-1}, PAP^{-1}, and PB_iP^{-1} take the form

$$CP^{-1} = [C,\ 0]$$

$$PAP^{-1} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \quad ; \quad PB_iP^{-1} = \begin{bmatrix} B^i_{11} & 0 \\ B^i_{21} & B^i_{22} \end{bmatrix}$$

Proof: Clearly if such a P exists then the system is not observable since $x_o = (0,\tilde{x})$ implies $y = 0$.

On the other hand, if there exists two indistinguishable states then there is a hyperplane of indistinguishable states. Hence

$$C\Phi_{(A+\Sigma u_i B_i)}^{\mathscr{K}} = 0$$

for some subspace $\mathscr{K}$. Let x be in $\mathscr{K}$ then x belongs to the kernel of C. Thus we may characterize $\mathscr{K}$ as the largest subspace of the kernel of C which is invariant under the action of* $\Phi_{(A+\Sigma u_i B_i)}$. If such a subspace exists then there exists a choice of basis such that $(A,\{B_i\},C)$ has the form indicated. □

The remark following Theorem 4 is relevant here as well.

We now give the observability version of Theorem 5.

Theorem 7: Any input-output map which can be realized by a bilinear system can be realized by one for which there are no indistinguishable states.

Proof: Use Theorem 6. If there are indistinguishable states then triangularize the system and delete the lower part of the systems. If the resulting system has indistinguishable states repeat the operation until there are no more indistinguishable states. □

Example: We can apply these results to a linear system with a linear or power law output. The n-dimensional scalar input, scalar output system

$$\dot{x} = Ax + bu \quad ; \quad y = (cx)^2 \quad ; \quad x(0) = 0$$

takes the form

$$\frac{d}{dt}\begin{bmatrix}1\\x\\x^{[2]}\end{bmatrix} = \begin{bmatrix}0 & 0 & 0\\ub & A & 0\\0 & uB & A^{[2]}\end{bmatrix}\begin{bmatrix}1\\x\\x^{[2]}\end{bmatrix} \quad ; \quad y = [0,\ 0,\ \tilde{C}]\begin{bmatrix}1\\x\\x^{[2]}\end{bmatrix} \quad (*)$$

Now if $(b, Ab, \ldots, A^{n-1}b)$ is of rank $n = \dim x$, then there is no vector space which contains the reachable set for the realization (*). The observability criterion can be applied to show that this system has no distinct indistinguishable states if $c; cA; \ldots cA^{n-1})$ is of rank n and $cA^i b$ is nonzero for some i.

*Here and above Φ with a subscript refers to a transition matrix associated with a linear system. See [9] section 4.

Equivalent Realizations

The state space isomorphism theorems for automata and linear systems are well known and of basic importance in these fields. Recently theorems of this type have appeared in other settings, for example [2] and [10]. Here we want to describe such a result for bilinear systems.

In this section we show that any two bilinear realizations of the same input-output may differ at most by a change of basis provided some natural minimality conditions are satisfied.

Let us agree to call x_o an <u>equilibrium</u> state of the bilinear system

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t) \ ; \ y(t) = Cx(t) \tag{I}$$

if Ax_o vanishes. This is the same as asking that x_o be an equilibrium solution of the differential equation which results when all the u_i are set to zero.

<u>Theorem</u> : Suppose that we are given two realizations of the same input-output map

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t);y(t) = Cx(t) \ ; \ x(0) = x_o$$

$$\dot{z}(t) = (F+\sum_{i=1}^{m} u_i(t)G_i)z(t);y(t) = Hz(t) \ ; \ z(0) = z_o$$

Let x_o and z_o be equilibrium states. Suppose that both systems are observable in that any two starting states can be distinguished for a suitable choice of u and suppose that the systems are controllable in that the reachable set from x_o or z_o is not confined to any proper linear subspace. Then the two realizations are equivalent.

<u>Proof:</u> Let the z-system be of dimension n. Without loss of generality we can assume the x-system is of dimension less than or equal to n. Let $u^1,u^2,\ldots,u^n$ be controls which are defined over the intervals $[0,t_1]$, $[0,t_2],\ldots[0,t_n]$ which result in z-trajectories $z^1,z^2,\ldots,z^n$. Let t_* be the largest of the t's and define $\tilde{u}^1,\tilde{u}^2,\ldots,\tilde{u}^n$ on $[0,t_*]$ by shifting the u^i to the latter portion of the interval and filling in

on the first portion with 0.

$$\tilde{u}^i(t) = \begin{cases} 0 & 0 \leq t \leq t_* - t_i \\ u^i(t-t_*+t_i) & t_*-t_i \leq t \leq t_* \end{cases}$$

Let $\tilde{z}^i$ be the resulting trajectory in the z system. As a result of the assumption that z_o is an equilibrium state

$$\tilde{z}^i(t) = \begin{cases} z_o & 0 \leq t \leq t_*-t_i \\ z^i(t-t_*+t_i) \quad ; & t_*-t_i \leq t \leq t_* \end{cases}$$

Let x^i be the trajectory which the x-systems generates under the control u^i. Because both systems generate the same input-output map we have

$$C\Phi_{(A+\Sigma u_iB_i)}(\tilde{x}^1(t_*),\tilde{x}^2(t_*),\ldots\tilde{x}^n(t_*))$$
$$= H\Phi_{(F+\Sigma u_iG_i}(\tilde{z}^1(t_*),\tilde{z}^2(t_*),\ldots\tilde{z}^n(t_*))$$

where $\Phi_{(A+\Sigma u_iB_i)}$ and $\Phi_{(F+\Sigma u_iG_i)}$ are the transition matrices which result from an arbitrary control u.

Now the matrix $Z = (\tilde{z}^1(t_*),\tilde{z}^2(t_*),\ldots,\tilde{z}^n(t_*))$ is non-singular by construction. If x-system is not of the same dimension as the z-system, or if the matrix $X = (\tilde{x}^1(t_*), \tilde{x}^2(t_*),\ldots,\tilde{x}^n(t_*))$ is singular then there exists a nonzero vector η such that $X\eta = 0$.

$$C\Phi_{(A+\Sigma u_iB_i)}X\eta = H\Phi_{(F+\Sigma u_iG_i)}Z\eta = 0$$

Thus $Z\eta$ is a starting state for the Z-system which is nonzero but equivalent to 0. This violates the observability hypothesis. Thus X must be a square matrix which is non-singular.

Since we have for all u

$$C\Phi_{(A+\Sigma u_iB_i)}X = H\Phi_{(F+\Sigma u_iG_i)}Z$$

and since I is certainly a possible transition matrix

$$CXZ^{-1} = H$$

Moreover, since no two states give rise to the same input-output map, the equality

$$C\Phi_{(A+\Sigma u_i B_i)} = CXZ^{-1}\Phi_{(F+\Sigma u_i G_i)}ZX^{-1}$$

implies

$$\Phi_{(A+\Sigma u_i B_i)} = XZ^{-1}\Phi_{(F+\Sigma u_i G_i)}ZX^{-1}$$

From this it follows that for $P = XZ^{-1}$

$$A = PFP^{-1}$$

$$B_i = PG_iP^{-1}$$

and from above

$$C = HP^{-1}$$

□

This result can also be used to establish isomorphism theorems for realizations in inhomogeneous form. That is, two realizations of the form

$$\dot{x}(t)=(A+\Sigma u_i(t)B_i)x(t)+\sum_{i=1}^{m} u_i(t)b_i \qquad y(t) = Cx(t)$$

can be shown to differ only by a choice of basis provided the appropriate minimality conditions are satisfied.

We point out that in actually determining equivalent realizations for systems and in the classification of systems, the results available in the study of Lie algebras (e.g. [4]) are of fundamental importance. Some recent work relating Lie algebras and system theoretic ideas is reported in [11].

Conclusions

In this paper we have shown that a particular bilinear model is both quite general and easy to work with. Building on previous results we have shown how to get a basic structure theory. There are many more specific problems

which can be examined using these tools. Some of these are under investigation and will be reported on soon.

References

1. R.E. Rink and R.R. Mohler, "Completely Controllable Bi-Linear Systems," SIAM J. Control, Vol. 6, No. 3, 1968.
2. R.W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM J. on Control, May 1972.
3. R.W. Brockett, "Algebraic Decomposition Methods for Non-linear Systems," in System Structure, IEEE Special Publication No. 71C61-CSS, August 1971 (S. Morse, Editor).
4. H. Samelson, Notes on Lie Algebras, Van Nostrand Co. 1969.
5. C. Lobry, "Controllabilite des systems non linearies," SIAM J. Control, Vol. 8, (1970), 573-605.
6. G.W. Haynes and H. Hermes, "Nonlinear Controllability via Lie Theory," SIAM J. Control, Vol. 8, (1970).
7. H. Sussmann and V. Jurdjevic, "Controllability of Non-linear Systems," J. of Differential Equations, to appear.
8. V. Jurdjevic and H.J. Sussmann, "Control Systems on Lie Groups," J. Differential Equations, to appear.
9. R.W. Brockett, Finite Dimensional Linear Systems, John Wiley and Sons, N.Y. 1970.
10. R.W. Brockett and A. Willsky, "Finite State Group Homomorphic Sequential Systems," IEEE Trans. on Automatic Control, August 1972.
11. R.W. Brockett and A. Rahimi, "Lie Algebras and Linear Differential Equations," in Ordinary Differential Equations, Academic Press, 1972. (L. Weiss, Editor).

LIE ALGEBRAS AND LIE GROUPS IN CONTROL THEORY

Roger W. Brockett

Harvard University

Cambridge, Massachusetts

---

*This work was supported in part by the U.S. Office of Naval Research under the Joint Services Electronics Program by Contract N00014-67-A-0298-0006 and by the National Aeronautics and Space Administration under Grant NGR 22-007-172.

PREFACE

The theory of differential equations and control have been linked very closely because most of the early applications of control theory were to engineering problems of the type which are most naturally described by ordinary differential equations. The questions of importance in control have helped to revitalize certain problem areas in differential equations and methods and tools from control have been useful in obtaining new results in differential equation theory. On the other hand, going back to the era of Lie himself, there has been close ties between Lie theory and differential equations. Thus it is not surprising that one finds that Lie theory and control are also closely connected. This "triangle" is the subject of this set of notes.

In control theory, Lie algebras make their appearance as Lie algebras of vector fields. Topological properties associated with Lie groups show up in the study of controllability and stability. Partial differential operators arise in the Fokker-Planck equations modeling the uncertainty of the environment and our uncertainty about the measurements we make of it. The problems which are of interest in control frequently require a generalization of the usual treatment of topics such as existence of geodesics, expressions for the spectrum of the Laplacian etc. The modification is, roughly speaking, to include the possibility of a metric which is "infinite" in certain directions, subject only to the condition that the directions along which it is finite can be combined in such a way as to make the distance between any two points finite. These notes contain a brief account of some of these topics, together with references where complete proofs can be found.

I have included a few exercises for the reader, both to indicate some results which do not exactly fit the format chosen here and to indicate some partial results and suggestions on additional problems of interest. Most of the examples are to be found in the exercises as well.

It is a pleasure to thank Prof. David Mayne for organizing such a stimulating forum for the exchange of ideas on system theory.

I. THE ALGEBRAIC THEORY OF LINEAR DIFFERENTIAL EQUATIONS

1.1 Lie Algebras and Linear Differential Equations

Clearly any linear differential equation of the form

$$\dot{x}(t) = A(t)x(t); \quad x(t) \in \mathbb{R}^n$$

can be expressed as

$$\dot{x}(t) = \left(\sum_{i=1}^{m} u_i(t)A_i\right)x(t)$$

with the A_i constant matrices and the $u_i(t)$ scalar functions of time. In view of the fact that the solution of the equation with a single A_i, i.e.

$$\dot{x}(t) = u(t)Ax(t)$$

is

$$x(t) = e^{A\int_0^t u(\sigma)\,d\sigma}\, x(0)$$

the question arises as to when the solution of the general problem can be written as the composition of a number of such solutions

$$x(t) = e^{A_1 g_1(t)}\, e^{A_2 g_2(t)} \cdots e^{A_m g_m(t)}\, x(0)$$

for a suitable choice of the $g_i(\cdot)$. Otherwise stated, we would like to know if the solutions of the matrix differential equation

$$\dot{X}(t) = (\sum_{i=1}^{m} u_i(t)A_i)X(t); \quad X(0) = I \quad \text{(identity)}$$

can be written as

$$X(t) = e^{A_1 g_1(t)}\, e^{A_2 g_2(t)} \cdots e^{A_m g_m(t)}$$

for a reasonably wide class of $u_i(t)$ and over some interval of time, say $|t| < \varepsilon$.

The above question is basically answered by a classical theorem of Frobenius [1]. However the theorem of Frobenius which applied here is a theorem in differential geometry. To use the insight of his result we need to look at the problem posed from a geometrical point of view. Consider the identity matrix as a point in the set of all nonsingular n by n matrices. Suppose that the one parameter curves $e^{A_i t}$ leave the identity as indicated in figure 1.

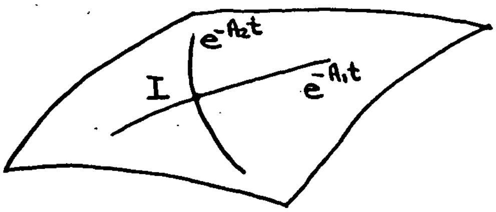

Figure 1: Neighborhood of I in the set of all n by n matrices

We regard the set of all points of the form

$$S = \{X : X = \prod_{i=1}^{m} e^{A_i \alpha_i}; \qquad \alpha_i \in \mathbb{R}^1\}$$

as a subset of the set of all nonsingular n by n matrices. Our question is, when do the integral curves of the given matrix differential equation corresponding to a wide class of $u_i(\cdot)$ lie in S? In order for this to be true for all piecewise continuous u's we require, for example, that

$$e^{A_1 t}\, e^{A_2 t}\, e^{-A_1 t}\, e^{-A_2 t}$$

be expressible as an element of S. To see why this is so we point out that the choice

$$u_1(\sigma) = \begin{cases} -1 & t \leqslant \sigma < 2t \\ 0 & 0 \leqslant \sigma < t; \quad 2t \leqslant \sigma < 3t \\ 1 & 3t \leqslant \sigma < 4t \end{cases}$$

$$u_2(\sigma) = \begin{cases} -1 & 0 \leqslant \sigma < t \\ 0 & t \leqslant \sigma < 2t; \quad 3t \leqslant \sigma < 4t \\ 1 & 2t \leqslant \sigma < 3t \end{cases}$$

$$u_i(\sigma) = 0 \quad i > 2$$

yields

$$X(4t) = e^{A_1 t} e^{A_2 t} e^{-A_1 t} e^{-A_2 t}$$

Geometrically, what we are asking is that in following the 4-sided path shown in figure 2 we should not be lead out of the set S.

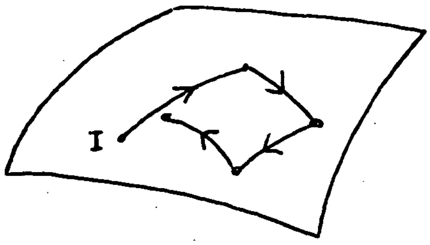

Figure 2: Illustrating the path leading to $e^{A_1 t} e^{A_2 t} e^{-A_1 t} e^{-A_2 t}$

More generally if f_1 and f_2 are smooth maps of $\mathbb{R}^n$ into $\mathbb{R}^n$ and if we apply the above choice of $u(\cdot)$ to the system

$$\dot{x}(t) = u_1(t)f[x(t)]+u_2(t)g[x(t)]; \quad x(0) = x_o$$

then a slightly messy calculation shows that to second order in t we have

$$x(4t) \simeq x_o + \{(\frac{\partial f}{\partial x})_{x=x_o} g(x_o) - (\frac{\partial g}{\partial x})_{x=x_o} f(x_o)\}t^2$$

The quantity $\frac{\partial f}{\partial x} g(x) - \frac{\partial g}{\partial x} f(x)$ is usually written as $[f,g]$ and is called the Lie bracket of f and g. One calls a set of vectors $f_i : \mathbb{R}^n \to \mathbb{R}^n$ <u>involutive</u> if the Lie bracket of any two is a linear combination of the $\{f_i\}$. Frobenius showed that the set of points near x_o which can be reached from x_o along integral curves of

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t) f_i(x)$$

with $\{f_i\}$ involutive can be expressed as

$$\phi_m(t_m, \ldots, \phi_3(t_3, \phi_2(t_2, \phi_1, x))) \ldots)$$

where $\phi_i(t,x)$ are the solutions of

$$\dot{x}(t) = f_i[x(t)]$$

The reason the set $\{f_i\}$ must be involutive is that otherwise the special choice of $u(\cdot)$ outlined above will, for small t, surely lead out of set of points expressible as $\phi_m(t_m,\phi(t_{m-1}, \ldots\phi(t,x_o))\ldots))$.

Applying this type of thinking to the linear case, we see first of all that the Lie bracket of A_1x and A_2x is $[A_1x, A_2x]=(A_1A_2-A_2A_1)x$ That is, the Lie bracket of the vector fields is expressible as the commutator of the matrices. We write $[A_i,A_j]$ for $A_iA_j-A_jA_i$. Thus if the set of matrices $\{A_i\}$ have the property that

$$[A_i,A_j] = \sum_{k=1}^{m} \gamma_{ijk}A_k$$

then the theorem of Frobenius would imply that for small $|t|$ we can write

$$X(t)x_o = \prod_{i=1}^{m} e^{A_ig_i(t)} x_o$$

A linear space of square matrices which is closed under $[\cdot,\cdot]$ is a matrix Lie algebra. Of course if the original set $\{A_i\}$ does not form a basis for a Lie algebra we simply supplement it with additional A's until it does. If x is of dimension n then there are only n^2 linearly independent matrices so this process always results in a finite set.

Wei and Norman [2] have given a direct verification of the above representation based on the implicit function theorem and have developed a set of nonlinear differential equations for the $g_i(\cdot)$. The basis for their derivation is the Baker-Campbell-Hausdorff formula

$$e^ABe^{-A} = B+[A,B]+ \frac{1}{2} [A,[A,B]+ \frac{1}{3!} [A,[A,[A,B]]\ldots]$$

Thus if one assumes a solution of the form

$$X(t)X_o= e^{A_1g_1(t)} e^{A_2g_2(t)} \ldots e^{A_mg_m(t)}$$

and then differentiates, the result is

$$\begin{aligned}\dot{X}(t) &= A_1\dot{g}_1(t)e^{A_1g_1(t)} e^{A_2g_2(t)} \ldots e^{A_mg_m(t)}\\ &+ e^{A_1g_1(t)} A_2\dot{g}_2(t)e^{A_2g_2(t)} \ldots e^{A_mg_m(t)}\\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\\ &+ e^{A_1g_1(t)} e^{A_2g_2(t)} \ldots A_m\dot{g}_m(t)e^{A_mg_m(t)}\end{aligned}$$

Now we must collect all the A's together at the left in order to compare this expression for $\dot{X}$ with that given by the differential equation. The Baker-Campbell-Hausdorff formula provides the means to do this. To see how this happens, observe that by inserting $e^{-A_ig_i(t)} e^{A_ig_i(t)}$ freely we can arrive at

$$\dot{g}_1A_1+\dot{g}_2e^{A_1g_1(t)}A_2e^{-A_1g_1(t)}+\ldots\dot{g}_me^{A_1g_1(t)}e^{A_2g_2(t)}\ldots A_m\ldots e^{A_1g_1(t)}$$

$$= A_1u_1(t)+A_2u_2(t) + \ldots + A_mu_m(t)$$

We apply the Baker-Campbell-Hausdorff expansion to each term on the left. If the set $\{A_i\}$ is a basis for a Lie algebra then we can express the result as a linear combination of the A_i. Since the A_i are linearly independent we can equate coefficients on each side and thereby get a set of differential equations for the g_i. It is important to note that the differential equations for the g_i only depend on the A_i through the commutation rules

$$[A_i,A_j] = \sum_{k=1}^{m} \gamma_{ijk}A_k$$

Thus when a differential equation is solved by this method a whole class of differential equations are solved at the same time -- one for each set of A's which satisfy the given commutation relation.

Exercises

1. Show that if the A_i in

$$\dot{X}(t) = \sum_{i=1}^{m} u_i(t)A_iX(t)$$

are all upper triangular then it is possible to express the solution of the differential equations for the $g_i(\cdot)$ explicitly in terms of integrals.

2. Show that the smallest Lie algebra of matrices which contains A_1 and A_2

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} ; \quad A_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

is 4 dimensional.

3. Study the definition of Euler angles from the point of view of the Wei-Norman equations. In particular explain why it is generally not possible to obtain a Wei-Norman representation the entire half-line $[0,\infty)$ in terms of the degeneracy of the Euler angles.

4. Show that for any square matrix P the set of all solutions of $PA+A'P = 0$ from a Lie algebra.

1.2 The $x^{[p]}$ and $x^{(p)}$ Equations

Associated with each linear map of $\mathbb{R}^n$ into $\mathbb{R}^n$ are two families of linear maps which may be described as follows. Choose a basis in $\mathbb{R}^n$ and let the original map be represented by the matrix A. Then we easily see that

$$y_i = \sum_{j=1}^{n} a_{ij}x_j$$

implies that the $n(n+1)/2$ linearly independent terms of the form y_iy_j depend linearly on the $n(n+1)/2$ linearly independent terms of the form x_ix_j. More generally the set of all linearly independent p-degree terms $y_iy_j\dots y_k$ depend linearly on the set of all linearly independent p-degree terms $x_ix_j\dots x_k$. How many linearly independent terms of degree p are there in n variables? If we denote this integer by N_n^p then it is easy to see that

$$N_{n+1}^{p+1} = N_{n+1}^{p} + N_n^{p+1}$$

from which an induction gives $N_n^p = \binom{n+p-1}{p}$. Thus associated with each map of $\mathbb{R}^n$ into $\mathbb{R}^n$ is a sequence of maps, the pth one mapping $\mathbb{R}^{N_n^p}$ into $\mathbb{R}^{N_n^p}$.

In order to give this family of maps a matrix description we need to choose a basis in $\mathbb{R}^{N_n^p}$ which is in some way convenient. The principle which guides our choice of basis is this: let $\langle x,y\rangle$ be the ordinary inner product

$$\langle x,y\rangle = \sum_{i=1}^{n} x_iy_i$$

If the map of $\mathbb{R}^n$ into $\mathbb{R}^n$ defined by A preserves length, we would like the maps of $\mathbb{R}^{N_n^p}$ into $\mathbb{R}^{N_n^p}$ to preserve length as well. To achieve this we introduce the basis elements

$$\sqrt{\binom{p}{p_1}\binom{p-p_1}{p_2}\cdots\binom{p-p_1-\cdots p_{p-1}}{p_p}}\; x_1^{p_1}x_2^{p_2}\cdots x_n^{p_n};\quad \sum_{i=1}^{n} p_i=p;\ p_i \geq 0$$

For example if $n=p=3$ we have basis elements

$$x_1^3,\ \sqrt{3}x_1^2x_2,\ \sqrt{3}x_1^2x_3,\ \sqrt{3}x_1x_2^2,\ \sqrt{6}x_1x_2x_3,\ \sqrt{3}x_1x_3^2,\ x_2^3,\ \sqrt{3}x_2^2x_3, \sqrt{3}x_2x_3^2, x_3^3$$

If we denote this vector, ordered lexigraphically, by $x^{[p]}$ then the choice of basis is such that $(||x||=(\langle x,x\rangle)^{1/2})$

$$||x^{[p]}|| = ||x||^p$$

More generally, we have

$$\langle x,y\rangle^p = \langle x^{[p]},y^{[p]}\rangle$$

We denote by $A^{[p]}$ the map, or matrix, which verifies

$$y = Ax \Rightarrow y^{[p]} = A^{[p]}x^{[p]}$$

The principle properties of $A^{[p]}$ are covered by the following theorem.

Theorem 1: Suppose we are given A and B. $A : \mathbb{R}^n \to \mathbb{R}^n$ and $B : \mathbb{R}^n \to \mathbb{R}^n$. Then $A^{[p]}$ and $B^{[p]}$ satisfy

i) $I_n^{[p]} = I_{N_n^p}$

ii) $(AB)^{[p]} = A^{[p]}B^{[p]}$

iii) $(A^q)^{[p]} = (A^{[p]})^q$; q integer; A^q defined

iv) $(A')^{[p]} = (A^{[p]})'$

Proof: i) Clear from definition. ii) Let $z=Ay=ABx$. Then $z^{[p]}=A^{[p]}y^{[p]}=A^{[p]}B^{[p]}x^{[p]}=[AB]^p x^{[p]}$. iii) This follows from ii) on letting $B=A$ (or $B=A^{-1}$ if A is invertible) and using induction. iv) This follows from the identity $\langle x,y\rangle^p=\langle x^{[p]},y^{[p]}\rangle$ and $\langle x,Ay\rangle = \langle A'x,y\rangle$.

A second series of maps associated with A are the so called compounds of A which we write as $A^{(p)}$ and define in terms of matrices as

$$A^{(p)} = \begin{pmatrix} \text{matrix of all p by p minors} \\ \text{of A ordered lexographically} \end{pmatrix}$$

Since there are $\binom{n}{p}$ ways to select the rows and $\binom{n}{p}$ ways to select the columns in a p by p minor of an n by n matrix we see that $A^{(p)}$ is an $\binom{n}{p}$ by $\binom{n}{p}$ matrix. The following properties of $A^{(p)}$ are well known. See for example [2] or [3].

Theorem 2: Let A and B be given; $A: \mathbb{R}^n \to \mathbb{R}^n$ and $B: \mathbb{R}^n \to \mathbb{R}^n$. Then $A^{(p)}$ and $B^{(p)}$ for $0 \leq p \leq n$ maps $\mathbb{R}^{\binom{n}{p}}$ into $\mathbb{R}^{\binom{n}{p}}$ and

i) $I_n^{(p)} = I_{\binom{n}{p}}$

ii) $(AB)^{(p)} = A^{(p)}B^{(p)}$

iii) $(A^q)^{(p)} = (A^{(p)})^q$ q integer; A^q defined

iv) $(A')^{(p)} = (A^{(p)})'$

We have used two different points of view in defining $A^{[p]}$ and $A^{(p)}$. The construction of $A^{[p]}$ from A was described in terms of linear maps whereas in the definition of $A^{(p)}$ we used matrices exclusively. Alternative approaches are available which give $A^{(p)}$ a geometric meaning in terms of skew symmetric forms of degree p in n variables.

These two constructions are specializations of the tensor product in the following way. If $A: \mathbb{R}^n \to \mathbb{R}^n$ and $B: \mathbb{R}^n \to \mathbb{R}^n$ then we may identify the tensor product of $A\eta$ and $B\gamma$ with $A\eta(B\gamma)'$; i.e.

$$A\eta \otimes B\gamma = A\eta(B\gamma)' = A(\eta\gamma')B'$$

If we consider the linear map of the space of n by n matrices into itself defined by $L(Q)=AQB'$ then $L_A(Q)=AQA'$ when restricted to act

on symmetric matrices has $A^{[2]}$ as a matrix representation and when restricted to the complementary space of skew symmetric matrices, it has $A^{(2)}$ as its matrix representation. Thus if we let $\simeq$ indicate "similar to" then we have

$$A \otimes A = A(\cdot)A' \simeq \begin{bmatrix} A^{[2]} & 0 \\ 0 & A^{(2)} \end{bmatrix}$$

One can also see that $A \otimes A \otimes A$ "contains" $A^{[3]}$ and $A^{(3)}$ but there are more than 2 symmetry types for a 3 index tensor so that $A^{[3]} \oplus A^{(3)}$ is only part of $A \otimes A \otimes A$. (Check the dimensionality; $n(n+1)(n+2)/6$ and $n(n-1)(n-2)/6$ does not add up to n^3.)

Now consider a linear differential equation in $\mathbb{R}^n$

$$\dot{x}(t) = A(t)x(t)$$

Observe that

$$x^{[p]}(t+h) = (I+hA(t))^{[p]}x^{[p]}(t)+0(h^2)$$

so that

$$x^{[p]}(t+h)-x^{[p]}(t) = [(I-hA)(t))^{[p]}-I]x^{[p]}(t)+0(h^2)$$

Thus

$$\frac{d}{dt}x^{[p]}(t) = (\lim_{h\to 0} [(I-hA(t))^{[p]}-I])x^{[p]}(t)$$

(Note that the dimensions of the identity matrices in these equations are n and N_n^p respectively.) We define $A_{[p]}$ to be the coefficient matrix in this differential equation.

$$\frac{d}{dt}x^{[p]}(t) = A_{[p]}(t)x^{[p]}(t); \quad p=1,2,3,\ldots$$

Thus the set of all p-degrees forms in $\{x_1,x_2,\ldots,x_n\}$ satisfies a linear differential equation with a coefficient matrix which is easily derived from A.

Starting with a matrix equation

$$\dot{X}(t) = A(t)X(t)$$

we can make an analogous construction using compound matrices (round brackets). The estimate

$$X^{(p)}(t+h) = (I+hA(t))^{(p)}X^{(p)}(t) + 0(h^2)$$

leads to

$$\frac{d}{dt}X^{(p)}(t) = (\lim_{h\to 0} [(I+hA(t))^{(p)}-I])X^{(p)}(t)$$

which we write as

$$\frac{d}{dt}X^{(p)}(t) = A_{(p)}(t)X^{(p)}(t); \quad p=1,2,\ldots,n$$

The special case in which p=n is the basis for well known Able-Jacobi-Liouville formula obtained by integrating the scalar equation

$$\frac{d}{dt}(\det X) = (\operatorname{tr} A(t))\det X(t)$$

Thus we see that $A_{[p]}$ and $A_{(p)}$ are infinitesimal versions of

$A^{[p]}$ and $A^{(p)}$ respectively. As such, they depend linearly on the elements of A. This has some significant implications.

We also have the infinitesimal version of the tensor product reduction given above. It takes the form

$$A(\cdot)+(\cdot)A' \simeq I \otimes A + A \otimes I \simeq \begin{bmatrix} A_{[2]} & 0 \\ 0 & A_{(2)} \end{bmatrix}$$

There are important relationships between A, $A_{[p]}$ and $A_{(p)}$ which are more or less clear from derivation. First of all, if A has all distinct eigenvalues $\{\lambda_i\}$ then the solutions of $\dot{x}(t)=Ax(t)$ consists of a sum of terms of the form $\alpha_i e^{\lambda_i t}$. Thus $x^{[p]}$ consists of products, p at a time, of such terms

$$x^{[p]} = \Sigma\beta_{ij\ldots k} e^{(\lambda_i+\lambda_j+\ldots\lambda_k)t}$$

Thus the eigenvalues of the $\binom{n+p-1}{p}$ by $\binom{n+p-1}{p}$ matrix $A_{[p]}$ are the $\binom{n+p-1}{p}$ sums over distinct (unordered) index sets

$$\lambda_i+\lambda_j+\ldots\lambda_k; \quad \text{p terms}$$

The same is true for the case where A has eigenvalues of higher multiplicity. Similarly, the eigenvalues of $A_{(p)}$ consist of sums p at a time of the eigenvalues of A but in this case the indices i,j,..,k must all be distinct.

A second fact involves the transition matrix $\Phi_A(t)$ which satisfies

$$\dot{\Phi}(t) = A(t)\Phi(t); \quad \Phi(0) = I$$

By the above construction we see that

$$\Phi_{A_{[p]}}(t) = \Phi_A^{[p]}(t)$$

and

$$\Phi_{A_{(p)}}(t) = \Phi_A^{(p)}(t)$$

(Again, the last of these is the Able-Jacobi-Liouville formula if p=n.)

Finally, if $\{A_i\}$ is a basis for a Lie algebra and if

$$[A_i,A_j] = \sum_{k=1}^{m} \gamma_{ijk}A_k$$

then

$$[A_{i_{[p]}},A_{j_{[p]}}] = \sum_{i=1}^{m} \gamma_{ijk}A_{k_{[p]}}$$

That is, the $\{A_{i_{[p]}}\}$ form a Lie algebra with the same structural

constants. To see this we need to show that

$$[A,B]_{[p]} = [A_{[p]}, B_{[p]}]$$

but this can be seen from the approximations

$$e^{[A,B]_{[p]}t^2} = (e^{[A,B]t^2})^{[p]}$$
$$\approx e^{A_{[p]}t}e^{B_{[p]}t}e^{-A_{[p]}t}e^{-B_{[p]}t}$$
$$\approx e^{[A_{[p]},B_{[p]}]t^2}$$

where in all cases the approximations are valid up to and including terms of second order in t. Identical formulas hold with [p] replaced by (p).

This circle of ideas is of great importance in the theory of representations of Lie algebras; see [4] or [5]. However in control theory and differential equations there exist many problems where one can use these ideas, and other ideas from representation theory, to simplify calculations and to provide insight. A particular example is the study of the moment equations for stochastic differential equations. See, for example, reference [6].

Exercises

1. Show that

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k(t) & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ -\sqrt{2}k(t) & -1 & \sqrt{2} \\ 0 & -\sqrt{2}k(t) & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

are an A, $A_{[2]}$ pair.

2. Show that $A^{[p]}$ is orthogonal if A is orthogonal. What about $A^{(p)}$?

3. Describe in full the decomposition of $A \otimes A \otimes A$.

4. Give a definition of $A^{[p]}$ for which $z = Ax$ implies $z^{[p]} = A^{[p]}x^{[p]}$ but which does not require A to be square.

1.3 Matrix Lie Algebras and the Matrix Exponential

In section 1 we saw that the solution of the differential equation

$$\dot{x}(t) = (\sum_{i=1}^{m} u_i(t)A_i)x(t); \quad x(0) = x_0$$

could be expressed for small $|t|$ as

$$x(t) = e^{A_1g_1(t)}e^{A_2g_2(t)}\dots e^{A_mg_m(t)}x_0$$

provided the A_i form a basis for a Lie algebra. On the strength of the theorem of Frobenius, similar statements can be made for

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t) f_i[x(t)]; \quad x(0) = x_o$$

provided the set of vectors $\{f_i(\cdot)\}$ are involutive. There is a sort of converse question. If the set $\{A_i\}$ does not form the basis for a Lie algebra to what extent is it necessary to add elements to these sets in order to cover all possibilities? We know already that by adding enough elements to $\{A_i\}$ so as to obtain a basis for a Lie algebra we can be assured of a representation of the above form. However, it might happen that for

$$\dot{x}(t) = u_1(t)A_1x(t)+u_2(t)A_2x(t); \quad x(t) \in \mathbb{R}^n$$

the smallest Lie algebra which contains A_1 and A_2 is of dimension n^2. Are all of the n^2-2 elements which we add in order to get a Lie algebra really necessary?

In 1939 Chow [7] published a generalization of an earlier theorem of Caratheodory proving that if some regularity conditions hold, then along solution curves of

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t) f_i[x(t)]; \quad x_o = x(0)$$

one can reach the same points as one can along the solution curves of

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t) f_i[x(t)] + \sum_{i=1}^{\nu} v_i(t) g_i[x(t)]$$

where $g_i(x)$ are obtained as Lie brackets of the f_i, Lie brackets of these Lie brackets, etc. Thus on the basis of this "reachability" theorem of Chow we see that no matter how many elements we must add to get a basis for a Lie algebra, nothing short of the full set will suffice.

We formalize this discussion as follows. Let B denote any subspace of $g\ell(n)$. Let $\{B\}_A$ denote the smallest Lie algebra which contains B. Let C be any subset of $G\ell(n)$ and let $\{C\}_G$ denote the smallest group which contains C.

<u>Theorem 1</u>: With the above definitions

$$\{\exp B\}_G = \{\exp \{B\}_A\}_G$$

Perhaps the most elementary proof of this result appears in [8].

After sufficient insight is built up it is frequently possible to evaluate $\{\exp \{B\}_A\}_G$ by inspection. The insight comes from a handful of special cases and general formulas such as $\exp A_{[p]} = (\exp A)^{[p]}$. The notation for the principle special cases is this:

We take the field to be $\mathbb{R}$ and let $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$.

$$
\begin{aligned}
g\ell(n) &= \{X : X = n \text{ by } n \text{ matrices}\} \\
s\ell(n) &= \{X : X \in g\ell(n);\ \operatorname{tr} X = 0\} \\
so(n) &= \{X : X \in g\ell(n);\ X'+X = 0\} \\
sp(n) &= \{X : X \in g\ell(n);\ X'J+JX = 0\}
\end{aligned}
$$

Matrices satisfying the last condition are often called Hamiltonian because they take the form familiar in Hamiltonian mechanics

$$\begin{bmatrix} A & Q \\ R & -A' \end{bmatrix} ; \quad Q = Q'; \quad R = R'$$

It is very important to keep in mind that $J^2 = -I$ so that $J^{-1} = -J$.

Associated with each of these algebras is a multiplicative group of matrices which are defined in a corresponding way

$$
\begin{aligned}
G\ell(n) &= \{X : X \text{ is } n \text{ by } n \text{ matrix};\ \det X \neq 0\} \\
S\ell(n) &= \{X : X \in G\ell(n);\ \det X = 1\} \\
So(n) &= \{X : X \in G\ell(n);\ X'X = I\} \\
Sp(n) &= \{X : X \in G\ell(n);\ X'JX = J\}
\end{aligned}
$$

These groups are called the general linear group, the special linear group, the special orthogonal group and the symplectic group, respectively.

It is easy to verify that in any of these cases exp X belongs to a particular group if X belongs to the corresponding algebra. This corresponds to the following well known facts

i) exp M is nonsingular for all M

ii) $\det(\exp M) = \exp(\operatorname{tr} M) = 1$ if $\operatorname{tr} M = 0$

iii) exp A is orthogonal if A is skew symmetric since $(e^A)' = e^{A'} = e^{-A} = (e^A)^{-1}$ if $A = -A'$.

iv) exp A is symplectic if A is Hamiltonian since $e^{A'} J e^A = Je^{J'A'J} e^A = J$ if $A'J+JA = 0$.

Notice that the set of n by n symmetric matrices do not form a Lie algebra; alternatively, the nonsingular symmetric matrices do not form a group.

The implication for the study of differential equations is as follows. If X is an n by n matrix which satisfies the equation

$$\dot{X}(t) = A(t)X(t)$$

Then of course the fundamental solution $\Phi_n(t)$ is going to belong to the general linear group. But if A at all points in time belongs to one of the above subalgebras of $g\ell(n)$ then $\Phi_A(t)$ will belong to the corresponding subgroup of $G\ell(n)$. This group-algebra relationship provides qualitative information about the solution without actually solving the equations of motion.

To what extent are the above maps of the algebra into the group actually onto the group? It is well known that a real nonsingular matrix need not have a real logarithm. Thus as far as the real field is concerned, exp does not map $g\ell(n)$ onto $G\ell(n)$. However if

the field is either the reals or the complexes, then every matrix sufficiently close to the identity does have a logarithm in the appropriate field and it is easy to see that exp maps a neighborhood of zero in the algebra onto a neighborhood of the identity in the group in a one to one way.

Exercises

1. Consider the set of n by n matrices whose column sums are zero. Show that they form a Lie algebra. If we denote this algebra by L then characterize $\{\exp L\}_G$.

2. Let so(p,q) denote the set of matrices satisfying

$$A'\Sigma(p,q) + \Sigma(p,q)A = 0$$

where $\Sigma(p,q)$ is defined by

$$\Sigma(p,q) = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$$

Show that this set of matrices forms a Lie algebra and show that for all matrices M in $\exp\{so(p,q)\}$ we have

$$\Sigma(p,q) = M'\Sigma(p,q)M$$

These are often called the pseudo orthogonal groups since they preserve the pseudo length $x'\Sigma(p,q)x$.

1.4 Cones and Semigroups

A semigroup of real n by n matrices is simply a subset of the n by n matrices which is closed under matrix multiplication. A cone in a real vector space is a subset closed under addition and multiplication by positive real numbers. Consider a real Lie algebra L in the set of n by n matrices. Let K be a conical subset of L. In general K will not be closed under Lie bracketing but it could be. Let $\{\exp K\}_{SG}$ indicate the smallest semigroup which contains exp K. As we will see, a number of problems in control lead to the question of characterizing $\{\exp K\}_{SG}$ in terms of K. The connection between a Lie algebra and its corresponding Lie group suggests analogous relationships between cones in the algebra and semigroups in the corresponding group. This kind of relationship is illustrated in the following example.

Example: Let K be the cone in $g\ell(n)$ consisting of all n by n matrices A such that A'+A is nonnegative definite. Then $\{\exp K\}_{SG}$ includes all orthogonal matrices since all skew symmetric matrices belong to K. Moreover, all symmetric matrices with eigenvalues greater than or equal to one belong to $\{\exp K\}_{SG}$ by well known properties of the exponential map. Thus by appealing to the fact that any matrix can be written in polar form $M = \Theta R$ with Θ orthogonal and R positive definite we see that if for all vectors x of unit length $||Mx||^2 = ||\Theta Rx||^2 = ||Rx||^2 \geq 1$ then M belongs to $\{\exp K\}_{SG}$. It is easy to see that if $||Mx|| < 1$ for some x of

unit length then we can not express M in the required way thus this condition is necessary and sufficient. We conclude that the semigroup of "expansive" matrices is the exponential of the non-negative definite ones. Likewise, the semigroup of (nonsingular) "contractive" matrices is the exponential of the cone of non-positive definite matrices.

This example can be generalized somewhat to give a theorem with broader scope.

Theorem 1: Let K be as above and let L_P be the Lie algebra of matrices satisfying $A'P+PA = 0$ with $P'P = I$. Then $\{\exp K \cap L_P\}_{SG} = \{\exp K\}_{SG} \cap \{\exp L_P\}_G$ i.e. the expansive matrices in $\{\exp L_P\}_G$.

Proof: Given any orthogonal matrix P, the group of matrices satisfying $M'PM = P$ has the property that the polar representations of each element has both its factors in the group. That is, if $M = e^{\Omega}e^{R}$ with e^{Ω} orthogonal and e^{R} positive definite and symmetric, then $e^{\Omega'}Pe^{\Omega} = P$, $e^{R}Pe^{R} = P$. To prove this we note that if $e^{R}e^{\Omega'}Pe^{\Omega}e^{R\Omega} = P$ then $e^{R}e^{\Omega'} = Pe^{-R}P'Pe^{-\Omega}P'$. However the term of the right is a polar decomposition since $Pe^{-R}P'$ is symmetric and positive definite and $Pe^{-\Omega}P'$ is orthogonal. Thus by uniqueness of the polar decomposition we see that $e^{R} = Pe^{-R}P'$ and $e^{\Omega} = Pe^{\Omega}P'$ which shows that each factor belongs to the given group.

Now if M has the polar form $M = e^{\Omega}e^{R}$ and if M belongs to $\{\exp K\}_{SG} \cap \{\exp L_P\}_G$ then $R \geq 0$ and Ω and R belong to L_P. Thus Ω belongs to $L_P \cap K$ and so does R.

Typically the relationship between a cone in the Lie algebra and the semigroup which the exponential maps it into is very difficult to describe. One problem of this type which has been investigated extensively arises in probability theory. Let $x_o \in \mathbb{R}^n$ have nonnegative components which sum to one. Suppose that x(t) evolves in time according to

$$\dot{x}(t) = A(t)x(t); \quad x(0) = x_o$$

If $A(\cdot)$ has the two properties:

(i) the off-diagonal elements of A(t) are nonnegative for all t
(ii) the sums of the columns of A(t) are zero for all t,

then x(t) will have nonnegative components which sum to one for all $t \geq 0$. This is equivalent to saying that subject to the above restrictions on $A(\cdot)$ the solution of the matrix equation

$$\dot{X}(t) = A(t)X(t); \quad X(0) = I \qquad (*)$$

is a stochastic matrix; i.e. a matrix with nonnegative entries whose columns sum to 1. The imbedding problem [9] is that of determining which stochastic matrices ϕ can be reached from the identity along solutions of (*) given only that A(t) must satisfy (i) and (ii). Of course the set of matrices which satisfy (i) and (ii) form a cone and the set of reachable matrices form a semi-

group. It is not true however that for $n > 2$ this semigroup consists of all stochastic matrices.

In control applications there is particular interest in the case of cones of the form

$$K = \{X : X = \alpha A + \Sigma\beta_i B_i; \quad \alpha \geq 0; \ \beta_i \text{ unrestricted}\}$$

i.e. cones which are half spaces. The first point to make is that by virtue of theorem 3.1 we may as well assume that the B_i form a basis for a Lie algebra since by adding elements to $\{B_i\}$ to make the basis of the Lie algebra generated by $\{B_i\}$ we do not enlarge the reachable set. Moreover, it is also clear from theorem 3.1 that

$$\{\exp\{A,B_i\}_A\}_G \supseteq \{\exp K\}_{SG} \supseteq \{\exp\{B_i\}_A\}_G$$

It is more or less clear that if e^{At} is periodic then

$$\{\exp\{A,B_i\}_A\}_G = \{\exp K\}_{SG}$$

and Jurdjevic and Sussmann [10] have shown that this is also true if e^{At} is almost periodic.

It is also true that $Ad_A^k B_i$ belongs to the Lie algebra generated by the B_i's then

$$\exp K = e^{\alpha A}\{\exp\{B_i\}_A\}_G$$

For a proof and some generalizations see the thesis of Hirschorn [11].

Exercises

1. Calculate $\{\exp N\}_{SG}$ where N is the cone

$$N = \{X : X = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}; \quad X+X' < 0\}$$

2. It is well known that the elements of $\Phi_A(t)$ are nonnegative for all $t \geq 0$ if $A(t)$ itself as elements which are nonnegative off the diagonal -- the diagonals may have any sign. Give an example which shows that $\{\exp K\}_{SG}$ is not the entire semigroup of square matrices with nonnegative entries if K is the cone of A's described above. (Find a matrix with positive entires and negative determinant.)

3. Explore the relationship between #2 and the imbedding problem.

II. INPUT-OUTPUT SYSTEMS

In this chapter we consider input/output systems which can be represented by a pair of equations of the form

$$\dot{X}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)X(t); \quad y(t) = C(X(t)) \qquad (*)$$

Here X is an n by n matrix as are A and $B_1, B_2, \ldots, B_m$; the map C is subject to certain restrictions to be described later. The

differential equation is said to be of the "right invariant type" because a multiplication on the right by a fixed element of $G\ell(n)$ gives an equation

$$\dot{X}(t)M = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t)M$$

which is again of the same form and with the same coefficient matrices. This is to be contrasted with an equation such as

$$\dot{X}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t)+X(t)(D+\sum_{i=1}^{m} u_i(t)E_i)$$

which does not have this invariance property. The basic idea is to understand as well as possible the properties of input-output maps which can be represented by equation (*). We will study controllability, observability and state space isomorphism theorems.

2.1 Controllability

If u_i is an m-dimensional piecewise continuous function of time and if t_i is a nonnegative number, then we given the pairs (u_i,t_i) a semigroup structure by defining

$$(u_1,t_1) \circ (u_2,t_2) = (u_1|u_2,t_1+t_2)$$

whereby $u_1|u_2 = u_3$ we mean

$$u_3(t) = \begin{cases} u_1(t); & 0 \le t < t_1 \\ u_2(t-t_1); & t_1 \le t < t_2 \end{cases}$$

This is the concatenation semigroup with due regard for the domain of definition of the functions being concatenated. We denote it by U^m.

Consider the time invariant control system

$$\dot{x}(t) = f[x(t),u(t)] \quad ; \quad x(t) \in \mathbb{R}^n \qquad (**)$$

with f well enough behaved so as to guarantee the existence of a unique solution for each starting point $x_0 \in \mathbb{R}^n$ and each $(u,t) \in U^m$. Let T^n be the semigroup of one to one continuous maps of $\mathbb{R}^n$ into $\mathbb{R}^n$ with composition as the semigroup operation. Then the control system (**) defines a homomorphism of U^m into T^n. We denote this homomorphism by ϕ and, by analogy with automata theory, call the image of U^m under ϕ the Myhill semigroup of the system.

The main thing which is special about bilinear systems is that the Myhill semigroup is easily identified with a matrix semigroup. That is, if we have a system in $\mathbb{R}^n$

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)x(t)$$

then the matrix equation

$$\dot{X}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t); \quad X(0) = I$$

describes the relationship between U^m and T^n -- each matrix being associated with an element of T^n in the standard way

$$M \mapsto f(x) = Mx$$

If A is absent in the above equation then it is clear that the Myhill semigroup is actually a group since if $u(\cdot) \in U^m$ steers the system from I to M at time t_1 then $v(\cdot) \in U^m$ and defined by

$$v(t) = -u(t_1-t)$$

steers the system to M^{-1} at $t = t_1$.

Given an initial state x_o, the set of states reachable from x_o can be identified with the set of points which x_o is mapped into by the various elements of the Myhill semigroup. That is, the Myhill semigroup acts on the state space

$$S : \Sigma \to \Sigma$$

The reachable set from x_o is the "orbit" through x_o defined by this action.

We now give various examples of reachability theorems.

<u>Theorem 1</u>: There exists a control which steers the system

$$\dot{X}(t) = (\sum_{i=1}^{m} u_i(t)B_i)X(t)$$

from X_o to X_1 in time $t_1 > 0$ if and only if $X_1X_o^{-1}$ belongs to $\{\exp\{B_i^o\}_A\}_G$.

<u>Proof</u>: This is an immediate consequence of Theorem 1.3.1.

It is also easy to see that if A belongs to $\{B_i\}_A$ then the reachable set for

$$\dot{X}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t)$$

is just the same as it would be if A were absent.

Notice that the reachable set does not depend on t_1 as long as t_1 is positive. If A is absent and if one restricts the controls to be bounded, say $|u_i(t)| \leq 1$ then all points of the above form are reachable after a suitably long time but the time required will depend on the point to be reached.

A second result which we want to use in a moment is this.

<u>Theorem 2</u>: The reachable set at time t for

$$\dot{X}(t) = (\tilde{A}+\sum_{i=1}^{m} u_i(t)\tilde{B}_i)X(t); \quad X(0) = I$$

and

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \qquad \tilde{B}_i = \begin{bmatrix} 0 & B_i \\ 0 & 0 \end{bmatrix}$$

with A square is

$$R(t) = e^{At}\{\exp \mathrm{Ad}_{\tilde{A}}, \tilde{B}_i\}_A\}_G$$

Here $\{\mathrm{Ad}_{\tilde{A}}, \tilde{B}_i\}_A$ indicates the smallest Lie algebra which contains $\{\tilde{B}_i\}_A$ and is closed under the action of $\mathrm{Ad}_{\tilde{A}}$.

<u>Proof:</u> See reference [8], Theorem 7.

We can combine theorems 1 and 2 in an obvious way to get the following more general result.

<u>Theorem 3:</u> The reachable set at time t for

$$\dot{X}(t) = \tilde{A}x(t) + \sum_{i=1}^{m} u_i(t)\tilde{B}_i X(t) + \sum_{i=1}^{q} v_i(t)\tilde{C}_i X(t)$$

where

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}; \quad \tilde{B}_i = \begin{bmatrix} 0 & 0 \\ 0 & B_i \end{bmatrix} \ ; \quad \tilde{C}_i = \begin{bmatrix} 0 & C_i \\ 0 & 0 \end{bmatrix}$$

with A and B_i square is

$$R(t) = \exp At\{\exp\{\mathrm{Ad}_{\tilde{A}}, \tilde{B}_i, \tilde{C}_i\}_A\}_G$$

Finally, one can get additional results by using a nice lemma of Jurdjevic and Sussmann [10].

<u>Theorem 4:</u> The reachable set for the $\mathbb{R}^n$ system at time t starting from x=0 at t=0 and governed by

$$\dot{x}(t) = (A + \sum_{i=1}^{m} u_i(t)B_i)x(t) + \sum_{i=1}^{p} v_i(t)g_i; \quad x(t) \in \mathbb{R}^n$$

is the vector space generated by $\{L_i^k g_i\}$ where k indicates powers and L_i is a basis for the Lie algebra generated by $\{A, B_i\}$.

<u>Proof:</u> To begin we observe that if x_1 is reached at $t=t_1$ starting from x=0 at t=0 using the control (u,v) then the control $(u,\alpha v)$ steers the system to αx_1 at $t=t_1$. Also, we know that if we write the system as

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ 1 \end{bmatrix} = \left(\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + u_i(t)\begin{bmatrix} B_i & 0 \\ 0 & 0 \end{bmatrix} + v_i\begin{bmatrix} 0 & g_i \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} x(t) \\ 1 \end{bmatrix}$$

then the reachable set has a nonempty interior in

$$R = \{\exp\{\tilde{A}, \tilde{B}, \tilde{G}\}_{LA}\}_G \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

where

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \ ; \quad \tilde{B}_i = \begin{bmatrix} B_i & 0 \\ 0 & 0 \end{bmatrix} \ ; \quad \tilde{G}_i = \begin{bmatrix} 0 & g_i \\ 0 & 0 \end{bmatrix}$$

There exists a nonzero control of the form (0,v) which steers the system back to zero at time $t=t_1$ from 0 at t=0 -- use u=0 and invoke standard linear theory. According to lemma 6.1 of [10] we obtain on taking perturbations about this control an open set in R containing 0. Using the cone property mentioned in the first sentence we see that the reachable set is a vector space. Lie algebras tell us which one.

A particular problem in controllability theory which has received a good deal of attention is

$$\dot{x}(t) = Ax(t)+u(t)b\langle c,x(t)\rangle \quad ; \quad x(t) \in \mathbb{R}^n$$

where $u(\cdot)$ is a scalar, and b is a column vector. Of course the linear system

$$\dot{x}(t) = Ax(t)+bv(t)$$

is controllable in $\mathbb{R}^n$ if and only if $(b,Ab,..,A^{n-1}b)$ is of full rank. If the linear system is controllable it might be supposed that the bilinear one is also controllable since if v is a control which drives the state of the linear system from x_o to x_1 then the control

$$u(t) = v(t)/\langle c,x(t)\rangle$$

drives the bilinear system from x_o to x_1. This argument has the obvious fallacy that $\langle c,x(t)\rangle$ might vanish along the trajectory leaving u(t) undefined. In particular, if $x(0) = 0$ then of course x vanishes identically for all future time. Thus the most one could hope for is that any nonzero state could be steered to any nonzero state. It turns out that this is too much to hope for also. A simple pair of examples which illustrate that no amount of work can salvage this argument and which at the same time suggest the nature of the problem are these.

Consider the system

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + u(t)\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

which has the form

$$\dot{x}(t) = Ax(t) + u(t)b\langle c,x(t)\rangle$$

with [A,b,c] a minimal realization of $s/(s^2-1)$. However for any given x_o there exists x_1 such that x_1 is not reachable from x_o because regardless of k, the off-diagonal elements of $(A+k(t)bc')$ are always positive so that $\phi(t,t_o)$, the transition matrix, has all entries nonnegative for $t > t_o$. Thus if x(0) has nonnegative entries for all $t > 0$. This argument shows that the system is not controllable.

Consider the system

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + k(t)\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$$

which has the form $\dot{x}(t) = Ax(t)+k(t)bc'x(t)$ with [A,b,c] a minimal realization of $s/(s^2+1)$. In this case we see that the system is controllable on $R^2-\{0\}$. (See reference [12] for details.)

Exercises

1. Show that the Myhill semigroup for the linear system

$$\dot{x}(t) = Ax(t)+bu(t); \quad x(t) \in \mathbb{R}^n$$

can be identified with the multiplicative matrix semigroup

$$S = \{X : X = \begin{bmatrix} e^{At} & x \\ 0 & 1 \end{bmatrix} ; \quad t \geq 0; \ x \in \operatorname{span}(b, Ab, \ldots A^{n-1}b)\}$$

2. Consider a bilinear system

$$\dot{x}(t) = Ax(t) + u(t)Bx(t)$$

on $\mathbb{R}^n - \{0\}$. Is it true that if there exists any state x_o such that all points in $\mathbb{R}^n - \{0\}$ are reachable from x_o then all states have this property?

3. Consider the linear system

$$\dot{X}(t) = A_\ell X(t) + X(t)A_r + \sum_{i=1}^{m} u_i(t)B_i$$

Here $X(t)$ is an n by q matrix and A_ℓ and A_r are n by n and q by q respectively; the B_i are n by q. Show that the Myhill semigroup equation can be identified with

$$\frac{d}{dt}\begin{bmatrix} S_1(t) & S_3(t) \\ 0 & S_2(t) \end{bmatrix} = \left(\begin{bmatrix} A_\ell & 0 \\ 0 & A_r \end{bmatrix} + \sum_{i=1}^{m} u_i(t)\begin{bmatrix} 0 & B_i \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} S_1(t) & S_3(t) \\ 0 & S_2(t) \end{bmatrix}$$

Show that the reachable set at time t for the Myhill equation is

$$\exp At \cdot \exp\{Ad_A, \tilde{B}_i\}$$

2.2 Observability

We now consider systems with an output

$$\dot{X}(t) = (A(t) + \sum_{i=1}^{m} u_i(t)B_i(t))X(t); \quad y(t) = C(X(t)); \quad X(t) \in G\ell(n)$$

The exact nature of the output map is not essential. We give the output space no structure -- it is just a set. The critical assumption is that there should exist subgroups H_ℓ and H_r of $G\ell(n)$ such that $C(X_1) = C(X_2)$ if and only if

$$H_1X_1H_2 = X_2$$

for some H_1 in H_ℓ and some H_2 in H_r. Under this assumption $C(X)$ identifies X to within a multiplication on the left by an element of H_ℓ and a multiplication on the right with an element of H_r. We call systems of this form homogeneous.

In such a set up, the observation of y, even over a period of time, can at most determine X to within a right multiplication by an element of H_r. Thus we might as well regard the system as evolving on the coset space $G\ell(n)/H_r$. Whether or not the observation of y and the knowledge of u over the interval $[0,\infty)$ serves to identify uniquely an element of X/H_r as a starting state is then subject to investigation.

<u>Theorem 1</u>: Consider the above system with H_r and H_ℓ given. Let R denote the set of X's reachable from I. Suppose that R is a group.

Then two points X_1H_r and X_2H_r in $G\ell(n)/H_r$ give rise to the same input/output map if and only if for each R_1 in R there exists $H_1(R)$ in H_ℓ such that

$$R^{-1}H_1(R)RX_1H_r = X_2H_r$$

If we denote by P the subgroup

$$P = \{X : R^{-1}XR \in H_\ell; \quad \forall R \in S\}$$

then any two elements of the form X_1H_r and $P_1X_1H_r$ with P_1 in P are not distinguishable.

Proof: If X_1H_r and X_2H_r are to be indistinguishable as starting states we must have

$$H_\ell R_i X_1 H_r = H_\ell R_i X_2 H_r$$

for all R_i in R. Since H_ℓ and H_r are groups and since R is a subgroup of $G\ell(n)$, the above condition is equivalent to asking that for each R_i in R there exist $H_1(R)$ in H_ℓ such that

$$R_i^{-1}H_1(R_i)R_iX_1H_r = X_2H_r$$

The remainder of the conclusions are clear.

Exercises

1. Assuming that the evolution equations are of the form

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m} u_i(t)B_ix(t); \quad y(t) = H_\ell x(t)H_r$$

with

$$H_\ell = \{\exp\{C_i\}_A\}_G; \quad H_r = \{\exp\{D_i\}_A\}_G$$

give an observability condition in terms of Lie algebras. (See ref. [8] for some results along this line.)

2. Apply the results of problem 1 to the bilinear problem

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m} u_i(t)B_ix(t); \quad y(t) = c[x(t)]$$

by identifying $\mathbb{R}^n$ with the n dimensional affine group modulo $G\ell(n)$.

2.3 Isomorphic Systems

The two scalar realizations

$$\dot{x}(t) = x(t)+u(t)x(t); \quad y(t) = x^3(t); \quad x(0) = 1$$

and

$$\dot{z}(t) = 3z(t)+3u(t)z(t); \quad y(t) = z(t); \quad z(0) = 1$$

realize the same input-output map. They are each controllable on $(0,\infty)$ and any two reachable states are distinguishable. They are related by the automorphism of the multiplicative group $(0,\infty)$ defined by

$$z = x^3$$

Thus despite the apparent differences between these two realizations they are closely related. The following theorem describes a general result of this type.

Theorem 1: Consider the two homogeneous realizations of the same input-output map

$$\dot{X}(t) = (A+\sum_{i=1}^{m} u_i(t)B_i)X(t); \quad y(t) = c[X(t)]$$

$$\dot{Z}(t) = (F+\sum_{i=1}^{m} u_i(t)G_i)Z(t); \quad y(t) = h[Z(t)]$$

which evolve in $G\ell(n_1)$ and $G\ell(n_2)$ respectively and which have reachable sets from the identity, R and $\hat{R}$, which are groups. Suppose H_ℓ, H_r and $\hat{H}_\ell$, $\hat{H}_r$ are given subgroups of $G\ell(n_1)$ and $\hat{G\ell}(n_2)$ respectively such that c and h are one to one on $H_\ell RH_r$ and $\hat{H}_\ell \hat{R}\hat{H}_r$ and such that the systems are observable on RH_r and $\hat{R}\hat{H}_r$. Finally, suppose that there is no normal subgroup of R which has a nontrivial intersection with $R \cap H_r$ and the same for $\hat{R}$ and $\hat{H}_r$. Then there exists an isomorphism $\phi : R \to \hat{R}$ such that

$$\phi(e^{At}) = e^{Ft}; \quad \phi(e^{B_i t}) = e^{G_i t}$$

Proof: Suppose that there exists a control (u,T) in U^m which takes the first system from I to $D_1 \neq I$ and takes the second system from I to I. Let D denote the set of all such points. By virtue of the observability hypothesis we see that D is a subset of H_r and, in fact, a subgroup of H_r. Moreover it is easily seen to be a normal subgroup of R and hence of $R \cap H_r$. By hypothesis D is trivial. This implies that there is a one to one correspondence between points in $R \cap H_r$ and $\hat{R} \cap \hat{H}_r$ which is, in fact, a homomorphism.

We see that R and $\hat{R}$ are both homomorphic images of U^m. If a pre-image of R in U^m is in U_R then what is the image under the action of the second system of U_R? It is clearly $\hat{R}$ or else a subgroup of $\hat{R}$. If it is a subgroup then the subgroup must contain $R \cap \hat{H}_r$ but there is a one to one and onto correspondence between $\hat{R}/R \cap H_r$ and $\hat{R}/\hat{R} \cap \hat{H}_r$ and an isomorphism between $R \cap H_r$ and $\hat{R} \cap \hat{H}_r$ Using the properties of the system maps we see that the above map must be onto R and thus it establishes an isomorphism. The remaining claims then follow.

Exercises

1. Develop the Lie algebra analog of Theorem 1.

2. Apply the above results to bilinear systems of the form

$$\dot{x}(t) = Ax(t)+\sum_{i=1}^{m} u_i(t)B_i x(t); \quad y(t) = cx(t); \; x(0) = x_o$$

See P. d'Allessandro, A. Isidori and A. Ruberti [13] and Brockett [14].

III. OPTIMAL CONTROL

This chapter is quite brief due to the absence in the literature of results relating specifically to the Lie group case. We discuss only two problem areas -- the question of existence of optimal controls in the bang bang case and questions centering around minimum "energy" transfer.

3.1 Bang-Bang Theorems

It is well known that under very weak assumptions on the matrices $A(\cdot)$ and $B(\cdot)$ the linear system

$$\dot{x}(t) = A(t)x(t)+B(t)u(t);\ x(0) = \text{given}$$

with controls constrained by

$$|u_i(t)| = 1$$

has a set of reachable points at any time $t_1 > 0$ which is the same as the set of points reachable with the constraint relaxed to

$$|u_i(t)| \leq 1$$

This is called a "bang-bang theorem" because the controls u_i need only take on their extreme values and not intermediate ones. Some generalizations of this have been investigated by Krenner [15] and Sussmann [16]. We examine only an easy case here.

<u>Theorem 1</u>: Let X satisfy the differential equation in $G\ell(n)$

$$\dot{X}(t) = AX(t) + \left(\sum_{i=1}^{m} B_i u_i(t)\right)X(t)$$

Then if $[\mathrm{Ad}_A^k(B_i),B_j]$ is zero for all i and j and $k=0,1,\ldots n^2-1$ then the set of states reachable at time t for $|u_i(t)| = 1$ is the same as the set reachable for $|u_i(t)| \leq 1$.

<u>Proof</u>: In view of the commutativity condition we can express the solution of the given equation as

$$X(t) = e^{At} e^{\int_0^t \sum_{i=1}^{m} e^{-A\sigma}B_i e^{A\sigma} u_i(\sigma)\,d\sigma} X(0)$$

See [8] Theorem 7 for details. Now since the bang-bang theorem is valid for the linear system

$$\dot{F}(t) = \sum_{i=1}^{m} e^{-At}B_i e^{At} u_i(t)$$

and since $X(t) = e^{At}e^{F(t)}$ we see that it holds for the systems defined here as well.

<u>Exercises</u>

1. The solution of the scalar differential equation

$$\dot{x}(t) = u(t)x(t)+v(t)$$

is

$$x(t) = e^{\int_0^t u(\sigma)d\sigma} x(0) + \int_0^t e^{\int_\sigma^t u(\rho)d\rho} v(\sigma)d\sigma$$

Is the bang-bang theorem valid if we regard u and v as controls?

2. Is the bang-bang theorem valid for the pair of scalar equations

$$\dot{z}(t) = u(t)z(t)$$
$$\dot{x}(t) = (u(t)+v(t))x(t)$$

3. Show that the bang-bang theorem is valid for

$$\dot{x}(t) = u(t)x(t)$$
$$\dot{y}(t) = -y(t)+u(t)$$

Generalize this result.

3.2 Least Squares Theory

Under the assumption used in the previous section we can develop a satisfactory theory for minimizing

$$\eta = \int_0^t \sum_{i=1}^m u_i^2(t)dt$$

subject to the constraint that the system

$$\dot{X}(t) = (A + \sum_{i=1}^m u_i(t)B_i)X(t) \qquad (*)$$

should be transferred from the state X_o at $t=0$ to the state X_1 at $t=t_1$.

<u>Theorem 1</u>: Let $X(t)$ satisfy the $G\ell(n)$ equation (*). Suppose that $[Ad_A^k B_i, B_j] = 0$ for all i and j and $k=0,1,2,\ldots n^2-1$. Suppose that

$$X_1X_0^{-1} \ \varepsilon e^{At_1}\{\exp\{Ad_A, B_i\}_A\}_G$$

Then there exists a control $u(\cdot)$ which steers the system from X_o at $t=0$ to X_1 at $t=t_1$ and minimizes η. This control is the same as the control which steers the linear system

$$\dot{F}(t) = \sum_{i=1}^m e^{-At}B_ie^{At}u_i(t)$$

from 0 at $t=0$ to $\ell n(e^{-At_1}X_1X_o^{-1})$ at $t=t_1$ and minimizes η where ℓn denotes the real solution of

$$e^M = e^{-At_1}X_1X_o^{-1}$$

which results in the smallest value of η. The optimal control is of the form

$$u_i(t) = tr(M_ie^{-At}B_ie^{At})$$

for some constant matrices M_i.

Proof: As in the proof of the bang-bang theorem we see that

$$X(t) = e^{At}e^{F(t)}$$

where F(t) satisfies

$$\dot{F}(t) = \sum_{i=1}^{m} e^{-At}B_i e^{At}u_i(t)$$

From this point on everything follows from standard linear theory. See [17], section 22.

Exercises

1. Consider the system

$$\dot{x}(t) = x(t)+u(t)$$
$$\dot{y}(t) = u(t)y(t)$$

Suppose we want to steer this system from (α,β) to (γ,δ) in t_1 units of time and to minimize

$$\eta = \int_0^{t_1} u^2(t)dt$$

If δ/β is positive this transfer is possible and the $u(\cdot)$ which achieves the optimal is of the form ae^t+b. Generalize Theorem 1 in such a way as to capture this example.

2. If B_1 and B_2 commute, describe the solutions of

$$\prod_{i=1}^{\nu} (e^{B_1u_i}e^{B_2v_i}) = N$$

IV. STOCHASTIC DIFFERENTIAL EQUATIONS

Stochastic processes on spheres has been of interest in physics for some time. Debye [18] in his book on statistical mechanics gives one application of S^2 stochastic processes. Nuclear magnetic resonance phenomena account for some more recent interest in diffusions on S^2. See Chapter 15 of the recent text [19]. The French mathematical physicist Perin wrote a classical paper [20] on diffusion on SO(3). Recent interest in physics regarding models of the type under study here is discussed in Fox [21]. Transmission of electromagnetic waves through random media leads to stochastic processes on the symplectic group -- distance playing the role usually assumed by time. Tutubalin [22] can be consulted for recent results and references. Carrier [23] has examined an equation of this general type in connection with a gravity wave propagation problem. One can think of this study as a stochastic process on the two dimensional symplectic group. An engineering problem for which the theory is potentially interesting is the randomly switched electrical circuit.

4.1 Bilinear Stochastic Equations

In this paper all stochastic differential equations are to be interpreted in the Ito sense. All Wiener processes are of unity variance and Wiener processes with distinct indices are assumed to be uncorrelated. The reader is encouraged to study Clark [24] for more details on stochastic calculus.

Under what circumstances does the Ito equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} dw_i(t)B_i x(t) \qquad (*)$$

evolve on the manifold defined by $x'Qx$ = constant? If we expand to second order keeping in mind that $dw_i dw_j = \frac{1}{2}\delta_{ij}dt$ we get

$$dx'Qx = x'(A'Q+QA)xdt + \sum_{i=1}^{m} x'(B_i'Q + QB_i)xdw + \frac{1}{2}\sum_{i=1}^{m} x'B_i'QB_i x\ dt$$

Thus in order for the derivative of $x'Qx$ to vanish we require

$$A'Q + QA + \frac{1}{2}\sum_{i=1}^{m} B_i'QB_i = 0$$

and also we require

$$B_i'Q + QB_i = 0$$

We see that the drift term A needs to be "corrected" by a term coming from the white noise. For example, if we want equation (*) to evolve on a sphere then A is not skew symmetric as it would be in the deterministic case but rather it has a correction term whose size depends on the B_i. On the other hand, the B_i must be skew symmetric.

In order to evolve on the symplectic group it is a skew symmetric form which must be preserved. Repeating the above with Hamiltonian matrices gives rise to the conditions that B_i and $A + \frac{1}{2}\sum_i B_i^2$ should be Hamiltonian.

<u>Exercises</u>

1. Show that the Ito equation

$$\begin{bmatrix} dx_1 & dx_2 \\ dx_3 & dx_4 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} dt + \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}\begin{bmatrix} dw_1 & dw_2 \\ dw_3 & -dw_1 \end{bmatrix}$$

evolves on the special linear group $S\ell(2)$ if suitable restrictions are placed on α, β, γ, δ.

2. Generalize the previous problem to $S\ell(n)$.

4.2 The Moment Equations

Associated with the stochastic equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} B_i x(t)dw_i(t) \qquad (*)$$

is a family of higher order equations analogous to those given in section 1.2. These are the equations for $x^{[p]}$. In order to display their form it is necessary to work out section 1.2 using the Ito calculus. As an alternative, suggested to me by Martin Clark, one can convert (*) into an analogous Stratonovich equation, use the ordinary calculus to get the $x^{[p]}$ equation, and then convert back to the Ito form. This idea is particularly attractive in the present setup since we have the deterministic results already.

The Stratonovich analog of (*) is simply

$$\hat{d}x(t) = (A-\frac{1}{2}\sum_{i=1}^{m} B_i^2)x(t)dt + \sum_{i=1}^{m} B_i x(t)\hat{d}w_i(t)$$

where $\hat{d}$ indicates Stratonovich differentials. Applying ordinary calculus we get

$$\hat{d}x^{[p]}(t) = (A-\frac{1}{2}\sum_{i=1}^{m} B_i^2)_{[p]}x^{[p]}dt+\sum_{i=1}^{m} B_i^{[p]}x^{[p]}(t)\hat{d}w_i(t)$$

Now if we want to convert this back to an Ito form we must correct the drift term to get

$$dx^{[p]}(t)=[(A-\frac{1}{2}\sum_{i=1}^{m} B_i^2)^{[p]}+\sum_{i=1}^{m}(B_i^{[p]})^2]x^{[p]}(t)dt+\sum_{i=1}^{m} B_i^{[p]}x^{[p]}(t)dw_i(t)$$

We can easily take expectations to get the moment equation

$$\frac{d}{dt}(\mathscr{E}x^{[p]}(t))=[(A-\frac{1}{2}\sum_{i=1}^{m} B_i^2)^{[p]}+\sum_{i=1}^{m}(B_i^{[p]})^2]\mathscr{E}x^{[p]}(t)$$

Notice that the apparently more general equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} B_i x(t)dw_i(t)+\sum_{i=1}^{m} e_i dw_2(t) \qquad (**)$$

is covered by these equations as well. To see this we let

$$\tilde{x} = \begin{bmatrix} x \\ 1 \end{bmatrix}$$

then x satisfies an equation of the form

$$d\tilde{x}(t) = \tilde{A}\tilde{x}(t)dt + \sum_{i=1}^{m}(\tilde{B}_i+\tilde{C}_i)\tilde{x}(t)dw_i(t)$$

There are many papers in the literature which analyze the stability of of these equations under various assumptions -- particular emphasis being placed on the case p=2. See, e.g. [25]. In reference [6] it is shown that under a suitable hypothesis all the moment equations are stable.

<u>Exercises</u>

1. Show that in the scalar case the moment equations for

$$dx(t) = \alpha(t)x(t)dt + \beta(t)x(t)dw(t)$$

are

$$\frac{d}{dt}\mathscr{E}x^p(t) = [p'\alpha(t)-\frac{1}{2}\beta^2(t))+\frac{1}{2}p^2\beta^2(t)]\mathscr{E}x^p(t)$$

Notice that if α and $\beta \neq 0$ are constant then it can never happen that all moment equations are stable.

2. A problem of interest in geophysics leads to the stochastic equation

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} 0 & dt \\ -dt+\varepsilon dw(t) & 0 \end{bmatrix} \begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} ; \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Show that the autocorrelation is, for small ϵ, approximated by

$$\mathscr{E}x_1(t)x_1(\tau) \approx e^{(\varepsilon^2/4)(t+\tau)}e^{-(\varepsilon^2/4)(t-\tau)}\cos(t-\tau)$$

(See Carrier [23]).

4.3 Fokker-Planck Equations

Associated with the Ito equation

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} dw_i B_i x(t)$$

is the formal Fokker-Plank equation

$$\frac{\partial \rho}{\partial t} - \frac{1}{2}\,\mathrm{tr}\Big(\sum_{i=1}^{m} B_i x x' B_i', \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\Big)\rho - \nabla_x \rho A x = 0$$

However, if x evolves on a manifold then this equation will not be especially useful unless the redundant variables are eliminated. In order to carry out this reduction it is necessary to coordinatize the manifold in some natural way. This coordinatization necessarily proceeds in a case by case way. To illustrate we work out four cases on the two-sphere S^2.

Consider the stochastic equations (Compare with McKean [26] who considers case b, case a being classical.)

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -dt & -dw_3 & dw_2 \\ dw_3 & -dt & -dw_1 \\ -dw_2 & dw_1 & -dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (a)$$

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}dt & 0 & dw_2 \\ 0 & -\frac{1}{2}dt & -dw_1 \\ -dw_2 & dw_1 & -dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (b)$$

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}dt & -dt & dw_2 \\ +dt & -\frac{1}{2}dt & -dw_1 \\ -dw_2 & dw_1 & -dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (c)$$

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} = \begin{bmatrix} 0 & -dt & 0 \\ dt & -\frac{1}{2}dt & -dw_1 \\ 0 & dw_1 & -\frac{1}{2}dt \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (d)$$

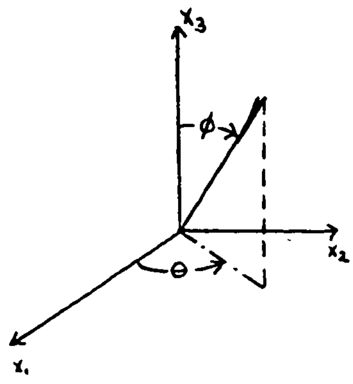

Figure 3:
Spherical Coordinates

We introduce polar coordinates according to figure 3. The Fokker-Planck equations corresponding to the above cases are then

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi} + \frac{1}{\sin^2\phi}\frac{\partial^2}{\partial\theta^2}\right)\right]\rho(t,\phi,\theta) = 0 \qquad \text{(a)}$$

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi} + \frac{1}{\tan^2\phi}\frac{\partial^2}{\partial\theta^2}\right)\right]\rho(t,\phi,\theta) = 0 \qquad \text{(b)}$$

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\sin\phi\frac{\partial}{\partial\phi} + \frac{1}{\tan^2\phi}\frac{\partial^2}{\partial\theta^2}\right) + \frac{\partial}{\partial\theta}\right]\rho(t,\phi,\theta) = 0 \qquad \text{(c)}$$

$$\left[\frac{\partial}{\partial t} - \frac{1}{2}\left(\sin\theta\frac{\partial}{\partial\phi} + \cot\phi\cos\theta\frac{\partial}{\partial\theta}\right)^2 + \frac{\partial}{\partial\theta}\right]\rho(t,\phi,\theta) = 0 \qquad \text{(d)}$$

The idea behind the derivation of these equations is that each of the three generators

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

can be associated with a first order partial differential operator which describes the effect of a drift around the corresponding axis of rotation and also with a second order partial differential operator which describes the effect of a diffusion around the corresponding axis of rotation. The derivation of these operators is an exercise in differential geometry, however the following insight is useful.

On a manifold with a Riemannian metric $(g_{ij}(x))$, the Laplace-Beltrami operator [27]

$$\nabla^2 = \frac{1}{\sqrt{\det(g_{ij}(x))}}\frac{\partial}{\partial x_i}(g_{ij}(x))^{-1}\sqrt{\det(g_{ij}(x))}\frac{\partial}{\partial x_j}$$

serves as the Laplacian, in that the basic heat equation, assuming constant conductivity, is

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\nabla^2\right)\phi(t,x) = 0$$

On S^2, in terms of the given coordinates, the usual metric is

$$(ds)^2 = [d\phi, \ d\theta]\begin{bmatrix} 1 & 0 \\ 0 & \sin^2\phi \end{bmatrix}\begin{bmatrix} d\phi \\ d\theta \end{bmatrix}$$

one sees easily that case a above corresponds to the heat equation.

As for case b, it is obtained from case a by removing one of the generators -- the one which corresponds to a diffusion about the x_3-axis. This is equivalent to subtracting $\frac{1}{2}(\partial^2/\partial\theta^2)$ from the operator appearing in case a.

Case c is obtained in an analogous way. We must add a drift term to the operator appearing in b corresponding to a rotation about the x_3-axis. Thus we add a $(\partial/\partial\theta)$ term to the operator appearing in b.

Case d is the most degenerate of all in that there is now only diffusion about one axis. There is a $(\partial/\partial\theta)$ drift term as in case c together with the operator which corresponds to diffusion about the x_1-axis.

It is of some interest to note that all these operators are studied in quantum theory. See Rose [28], appendix A.

Exercises

1. Consider the stochastic equation

$$\begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} \quad \begin{bmatrix} -\frac{1}{2}\,dt & -dw & 0 \\ dw & -\frac{1}{2}\,dt & dt \\ 0 & dt & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \; ; \; x_1^2(0)+x_2^2(0)-x_3^2(0)=1$$

Show that it evolves on the manifold defined by $x_1^2+x_2^2-x_3^2=1$. Introduce coordinates in this manifold and work out the Fokker-Planck equation. Is there a limiting distribution?

2. Show that the moment equations associated with each of the four cases analyzed here are stable. (see [26])

4.4 Calculation of Diffusion Times

We continue with the analysis of the four cases of diffusions on spheres, now with a view toward determining, if possible, a complete solution to the Fokker-Plank equation. In cases where that proves too difficult we look for some measure of the relaxation time of the process.

To begin with, the standard S^2 diffusion (case a above) leads to the Fokker-Plank equation

$$\frac{\partial\rho(t,x)}{\partial t} - \frac{1}{2}\,\nabla^2\rho(t,x) = 0$$

Where ∇^2 is the usual Laplacian on the sphere. It is, of course, well known that the eigenvalues of the Laplacian on the sphere are $n(n+1)$, $n=0,1,2,\ldots$ with the nth being of multiplicity $2n+1$. Thus the general solution of the above equation starting from the singular distribution concentrated at $\theta = \phi = 0$ is

$$\rho(t,\theta,\phi) = \sum_{n=1}^{\infty}\sum_{k=-n}^{n} P_{nk}(\cos\phi)e^{ik\theta}e^{-n(n+1)t}$$

where P_{nk} are the spherical harmonics. We also see that the eigenvalues are a measure of the speed with which the density approaches steady state.

On the basis of this Green's function one can, of course, express the general solution of the Fokker-Plank equation in terms of its initial value. Thus we have, in terms of the spherical harmonics, a complete solution to the Fokker-Planck equation. This is classical.

On the other hand, it is possible to be almost as explicit in the other cases as well. This comes about because the 2n+1 equations for the coefficients of the spherical harmonics of the form $P_{nk}(\cos\ )e^{ik\theta}$ $k=0,\pm 1,\ldots\pm n$ are decoupled from those corresponding to $P_{n'k}(\cos\ )e^{ik\theta}$ for $n\neq n'$. Thus the solution of the Fokker-Planck equation reduces to a sequence of linear differential equations; the nth entry in the sequence being a coupled set of 2n+1 equations. It happens, however, that there is a simple connection between the moment equations of section 4.2 and the equations for the coefficients of the spherical harmonics. We describe this for the S^2 situation but similar results hold on spheres of any dimension.

For an S^2 equation x is a 3-vector and $x^{[p]}$ is of dimension $(p+1)(p+2)/2$. The equation for $x^{[p]}$ includes all linearly independent p-forms in x; thus it includes $(p-1)(p)/2$ terms of the form

$$(x_1^2+x_2^2+x_3^2)x^{[p-2]}$$

Hence we can partition $x^{[p]}$ into two parts of dimension $(p-1)p/2$ and $(p+1)(p+2)/2 - (p-1)p/2 = 2p+1$, respectively according to whether the components have a factor of $x_1^2+x_2^2+x_3^2$ or not. Now of course $x_1^2+x_2^2+x_3^2 = 1$ so that the components which do contain this factor can be thought of as moment equations of a lower order and hence they evolve independently of the second part of the equation. On the other hand, the 2p+1 components which do not contain $x_1^2+x_2^2+x_3^2$ as a factor evolve independently as well. Collecting these facts we see that the moment equations have the structure

$$\frac{d}{dt}\mathcal{E}x^{[p]}(t)=\begin{bmatrix} A_\delta & 0 & & \cdots & \\ 0 & A_{\delta+2} & & \cdots & \\ \cdots & \cdots & \vdots & \cdots & \\ \cdots & \cdots & \tilde{A}_{p-2} & 0 \\ \cdots & \cdots & 0 & \tilde{A}_p \end{bmatrix}\mathcal{E}x^{[p]}(t)$$

where δ is zero or one depending on whether p is even or odd. The dimension of A_p is $(2p+1)$ by $(2p+1)$ and the coefficients of the spherical harmonics of type P_{nk}, n fixed, $k=0,\pm 1,\pm 2,\ldots\pm n$ are governed by the differential equation

$$\dot{y}(t) = \tilde{A}_p y(t)$$

Thus the spectrum of the operators

$$(A - \frac{1}{2}\sum_{i=1}^{m} B_i^2)^{[p]} + \sum_{i=1}^{m} (B_i^{[p]})^2$$

which were derived in section 4.2, governs the relaxation time of the process. In case a above we have already commented that the spectrum is $\frac{1}{2}(n(n+1))$ with the nth term being of multiplicity 2n+1. In case b there is less diffusion and one would expect the relaxation to be slower. This is the case; a calculation shows that the first few entries of the spectrum compares with case as follows.

$$\frac{1}{2}\left\{\begin{array}{llllllllll} 0, & 2, & 2, & 2 & 6, & 6, & 6, & 6 & 6 \ldots & \text{case a} \\ \underbrace{0,}_{\text{I}} & 1, & \underbrace{1, \quad 2}_{\text{II}} & & \underbrace{2, \quad 2, \quad 5, \quad 5,}_{\text{III}} & & & & 6 \ldots & \text{case b} \end{array}\right.$$

Finally, we remark that examples b, c, and d are specific cases of the hypoelliptic operators of Hormander [29].

Exercises

1. Consider the linear stochastic equation

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} -1/2 & 1 \\ -1 & -1/2 \end{bmatrix} dt + \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix}; \quad x(0) = 0$$

as an approximation to the first two components of the S^2 equation

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}dt & dt & dw_1 \\ -dt & -\frac{1}{2}dt & dw_2 \\ -dw_1 & -dw_2 & -dt \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}; \quad x(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Compute the second moment in each case and compare.

2. Consider the stochastic equation on S^2 defined by

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \end{bmatrix} = \begin{bmatrix} -dt/2 & dw_1 & 0 \\ -dw_1 & -(1+\rho)dt/2 & \rho dw_2 \\ 0 & -\rho dw_2 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

Find the first few eigenvalues of corresponding Fokker-Planck operator as a function of ρ.

V. STABILITY THEORY

In the study of ordinary differential equations on Lie groups both linear and nonlinear problems are of interest, however in these notes we discuss linear problems only. Of course the most common stability problems encountered in control concern the general linear group. However in the study of specific applications other groups may occur. For example, in the case of problems

arising in classical mechanics the symplectic group plays a major role. Moreover since tensoring will typically transform a system evolving in $G\ell(n)$ into one which evolves on some subgroup of $G\ell(q)$ is desirable to take a general point of view.

5.1 Stability of the $x^{[p]}$ Equations

The following theorem is an obvious consequence of the calculations in section 1.2.

Theorem 1: The null solution of the system

$$\dot{x}(t) = A(t)x(t)$$

is stable (asymptotically stable) if and only if the null solution of the equation

$$\dot{y}(t) = A_{[p]}(t)y(t)$$

is stable (asymptotically stable). Moreover if all solutions of the first equation are bounded by $|x(t)| \leq Me^{-\lambda t}$ then all solutions of the second are bounded by $|y(t)| \leq M_1 e^{-p\lambda t}$.

When combined with standard estimates this theorem can give very precise information about high order systems which are either in the form of $\dot{y}(t) = A_{[p]}(t)y(t)$ or else in the form

$$\dot{y}(t) = A_{[p]}(t)y(t) + D(t)y(t)$$

with $D(t)$ small in some sense.

Example: We know from Liapunov [see e.g. [30]] that all solutions of the $Sp(2)$ equation

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -p(t) & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

are bounded if $p(\cdot)$ is pointwise nonnegative, periodic of period T with positive average value and with

$$\int_0^T p(t)dt \leq 4/T$$

Thus we see that all solutions of the $x^{[2]}$ equation

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \dot{y}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -2p(t) & 0 & 2 \\ 0 & -p(t) & 0 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix}$$

are also bounded under the same hypothesis. (Here we have taken $y_2 = 2x_1^2$ instead of $\sqrt{2}\, x_1^2$.) A change of basis puts this equation in a more symmetric form

$$\begin{bmatrix}\dot z_1(t)\\ \dot z_2(t)\\ \dot z_3(t)\end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2}(1-p(t)) & 0\\ -\frac{1}{2}(1-p(t)) & 0 & \frac{1}{2}(1+p(t))\\ 0 & \frac{1}{2}(1+p(t)) & 0\end{bmatrix}\begin{bmatrix}z_1(t)\\ z_2(t)\\ z_3(t)\end{bmatrix}$$

This equation evolves on the pseudo-orthogonal group $SO(2,1)$.

One particular fact which should be mentioned here is that systems with a single time varying parameter, say

$$\dot x(t) = Ax(t)+k(t)Bx(t) \qquad (*)$$

go into systems with a single time varying parameter e.g.

$$\dot x^{[p]}(t) = (A_{[p]}+k(t)B_{[p]})x^{[p]}(t)$$

Thus the many useful results about (*) (circle criterion, [17], etc.) can be extended in a nontrivial way.

Exercises

1. It is known that all solutions of the differential equation

$$\ddot x + \dot x + k(t)x(t) = 0$$

remain bounded if $0 \leq k(t) \leq \sim 3.9$ (see [17]). On the other hand, if one picks a positive definite quadratic form in x and $\dot x$ say $v(x,\dot x)$ and computes its derivative along solutions of the given differential equation then there exists one quadratic form which implies stability via Liapunov theory, for $0 \leq k(t) \leq 1$ but the constant 1 cannot be improved on using a quadratic Liapunov function. However, if we look at the $x^{[p]}$ version of the differential equation then a quadratic Liapunov function for $x^{[p]}$ is a 2p-degree Liapunov function for the original equation and a more suitable Liapunov function can be found. Work out the details.

2. Consider a differential equation in $\mathbb{R}^n$

$$\dot x(t) = Ax(t)+k(t)Bx(t)$$

Suppose that A and B generate a four dimensional Lie algebra which is isomorphic with $g\ell(2)$. Use the theory of the representations of $g\ell(2)$ (see, e.g. Samelson [4] page 114) and the circle criterion (see, e.g. [17]) to derive stability criteria for the given system.

5.2 Periodic Self-Contragradient Systems

A matrix Lie algebra is said to be self-contragradient if there exists a matrix P such that

$$PLP^{-1} = -L'$$

for all L in the Lie algebra. For example, $so(n)$ is self-contragradient with $P=I$ and $sp(n)$ is self-contragradient with $P=J$. As far as the stability of periodic systems is concerned, the important consequence of this assumption is that if $A(t)$ satisfies $PA(t)P^{-1} = -A'(t)$ then the transition matrix for

$$\dot{x}(t) = A(t)x(t)$$

satisfies

$$\Phi_A'(t)P\Phi_A(t) = P$$

since

$$\frac{d}{dt}(\Phi_A'(t)P\Phi_A(t)) = \Phi_A'(t)(A'(t)P+PA(t))\Phi_A(t) = 0$$

Thus $\Phi_A(t)$ similar to $(\Phi_A^{-1})'$. As an immediate consequence of this fact we see that the eigenvalues of $\Phi_A(t)$ occur in reciprocal pairs -- if λ is an eigenvalue then so is $1/\lambda$. If we assume we are dealing with real systems then of course the eigenvalues occur in complex conjugate pairs as well.

If $A(t) = A(t+T)$ then the well known Floquet theory insures that the transition matrix for

$$\dot{x}(t) = A(t)x(t)$$

can be expressed as

$$\Phi_A(t) = Q(t)e^{Rt}; \quad Q(0) = I$$

with $Q(t+T) = Q(t)$ and R constant, though not necessarily real. The value of $\Phi_A(T)$ is decisive as far as the stability of a periodic system is concerned since $\Phi_A(nT) = [\Phi_A(T)]^n$.

If $A(t)$ is given by

$$A(t) = \sum_{i=1}^{m} a_i(t)A_i$$

with the A_i being a basis for a self-contragradient representation of a Lie algebra, then of course

$$\Phi_A'(t)P\Phi_A(t) = P$$

for all t. If $(\Phi_A(T))^n$ is bounded for $n=1,2,\ldots$ then we call $\Phi_A(T)$ stable. We call it P-strongly stable if it happens that for all sufficiently small R such that $R'P+PR = 0$, the matrix $e^R\Phi_A(T)$ is also stable. (Compare with [31].) In view of the fact that the eigenvalues of a matrix depend continuously on the elements of the matrix and in view of the fact that the eigenvalues of Φ_A must occur in reciprocal pairs, we see that if the eigenvalues of $\Phi_A(T)$ are distinct and if $\Phi_A(T)$ is stable, then it is P-strongly stable. However it can happen that $\Phi_A(T)$ is P-strongly stable even if the eigenvalues of $\Phi_A(T)$ are not distinct.

<u>Theorem 1</u>: If $\{A_i\}$ is the basis for a self-contragradient matrix Lie algebra, $A_i'P+PA_i = 0$, and if

$$\dot{x}(t) = (\sum_{i=1}^{m} a_i(t)A_i)x(t)$$

is periodic and if $\Phi_A(T)$ is P-strongly stable, then there exists $\varepsilon > 0$ such that for $|b_i(t)-a_i(t)| < \varepsilon$ and $b_i(t)$ periodic of period T the system

$$\dot{x}(t) = (\sum_{i=1}^{m} b_i(t)A_i)x(t)$$

is stable.

Exercises

1. Determine if for P = J the matrix

$$M = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & \cos & 0 & \sin\theta \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & -\sin & 0 & \cos\theta \end{bmatrix} \quad ; \; 0 < \theta < \pi$$

is P-strongly stable or not. See [30], theorem 8.

2. Show that if $p(t)$ is periodic of period T with average value zero and if

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -p(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

then $\Phi_A(T)$ is symplectic although $\Phi_A(t)$ for $t \neq T$ need not be. The corresponding $x^{[2]}$ equation is expressible as

$$\frac{d}{dt}\begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ -1 & -p(t) & -1 \\ 0 & -2 & -2p(t) \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$$

Use the idea of strong stability to investigate the stability of these systems.

3. If D is diagonal then D+H is similar to a diagonal matrix if H is any symmetric matrix. However if D is diagonal there may exist an $n(n-1)/2$ dimensional set, the upper triangular matrices, such that D+T is not diagonalizable; consider the identity. Relate this to strong stability.

5.3 The Symplectic Case

In the special case of the symplectic group Krein [30] has given an elegant theorem on how large the perturbation in Theorem 1 of the previous section can be. We give an application of this theorem and some facts about realizations of feedback systems as well.

Notice that the second order system with Q(t) symmetric

$$\ddot{x}(t) + Q(t)x(t) = 0; \quad x(t) \in \mathbb{R}^n$$

is equivalent to the symplectic system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -Q(t) & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Kreĭn has investigated this set of equations and more general ones. One of his results reads as follows.

Theorem 1: Let $P(t) = P(t+T) = P'(t)$, then all solutions of the equation

$$\ddot{x}(t)+P(t)x(t) = 0$$

are bounded if

i) $P(t) \geq 0$ all t

ii) $\int_0^T P(t)dt > 0$ (positive definite)

iii) $(4/T)I-\int_0^T P(t)dt > 0$ (positive definite)

Proof: See Krein [30], page 165.

As an example of an application of this result to problems of the type which arise frequently in system theory we prove the following theorem. (Compare with [32])

Theorem 2: Suppose that $q(s)$ and $p(s)$ are polynomials without common factors. Suppose further that $q(s)/p(s)$ is an even function of s with all its poles and zeros on the imaginary axis and assume that the poles and zeros of $sq(s)/p(s)$ interlace. Let $D = d/dt$ and let $k(\ )$ be periodic of period T. Then all solutions of the nth order differential equation

$$p(D)x(t) + k(t)q(D)x(t) = 0$$

are bounded provided

$$0 < \int |\lambda(t)|^2 dt < 4/T$$

where $\lambda(t)$ denotes the zero of $p(s)+k(t)q(s) = 0$ which has the largest magnitude.

Proof: Write $q(s)/p(s)$ as $r(s^2)/m(s^2)$ with r and m polynomials. This is possible because $q(s)/p(s)$ is even. Write $r(s)/m(s)$ as $b'(Is-A)^{-1}b$ with $A = A'$. This is possible because the poles and zeros of $r(s)/m(s)$ interlace, (See [25]). Thus

$$q(s)/p(s) = b'(Is^2-A)^{-1}b$$

and the differential equation in the theorem statement is equivalent to the system

$$\ddot{x} + (A+k(t)bb')x(t) = 0$$

Krein's result implies stability if

$$I(T/4) - \int_0^T (A+k(t)bb')dt > 0$$

But since the largest eigenvalue of the sum of two positive definite symmetric matrices is less than or equal to the sum of the largest eigenvalues of the respective matrices there is a corresponding inequality for integrals and we see that

$$\lambda_{max}\int_0^T (A+k(t)bb')dt \leq \int_0^1 |\lambda(t)|^2 dt$$

The result then follows.

It is interesting to compare this result with the analogous facts about completely symmetric systems investigated in [25]. Also notice that this theorem captures Liapunov's original theorem as a special case, as does the basic theorem of Krein.

Exercises

1. Use these results to investigate the stability of the scalar equation

$$x^{(4)}+4x^{(2)}+3x+k(t)(x^{(2)}+x) = 0$$

with k(t) periodic.

2. Derive a matrix version of Theorem 2.

VI. REFERENCES

1. H. Flanders, Differential Forms, Academic Press, 1963, New York.
2. J. Wei and E. Norman, "On the Global Representation of the Solutions of Linear Differential Equations as a Product of Exponentials," Proc. Am. Math. Soc., April 1964.
3. R. F. Gantmacher, Theory of Matrices, Chelsea, New York, 1959.
4. H. Samelson, Notes on Lie Algebras, Van Nostrand, New York, 1969.
5. M. Hausner and J. T. Schwartz, Lie Groups ; Lie Algebras, Gordon and Breach, London, 1968.
6. R. W. Brockett, "Lie Theory and Control Systems Defined on Spheres," SIAM J. on Applied Math., Vol. 24, No. 5, Sept. 1973.
7. W. L. Chow, "Über Systeme Von linearen partiellen Differentialgleichungen erster Ordinung," Math. Ann. Vol. 117, (1939), pp. 98-105.
8. R. W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM Journal on Control, Vol. 10, No. 2, May 1972, pp. 265-284.
9. S. Jonhansen (This Volume).
10. V. Jurdjevic and H. J. Sussmann, "Control Systems on Lie Groups," J. of Differential Equations, Vol. 12, No. 2, (1972) pp. 313-329.
11. R. Hirschorn, Ph.D. Thesis, Applied Mathematics, Harvard University, 1973.
12. R. E. Rink and R. R. Mohler, "Completely Controllable Bilinear Systems," SIAM J. on Control, Vol. 6, No. 3, 1968.
13. P. d'Allessandro, A. Isidori and A. Ruberti, "Realization and Structure Theory of Bilinear Dynamical Systems," (to appear, SIAM J. on Control).
14. R. W. Brockett, "On the Algebraic Structure of Bilinear Systems," in Variable Structure Control Systems, (R. Mohler and A. Ruberti, eds.) Academic Press, 1972, pp. 153-168.

15. A. J. Krener, A Generalization of the Pontryagin Maximal Principle and the Bang-Bang Principle, Ph.D. Thesis, Dept. of Mathematics, University of California, Berkeley, 1971.
16. H. J. Sussmann, "The Bang-Bang Problem for Linear Control Systems in $G\ell(n)$," SIAM J. on Control, Vol. 10, No. 3, Aug. 1972.
17. R. W. Brockett, Finite Dimensional Linear Systems, John Wiley and Sons, New York, 1970.
18. R. Debye, "Polar Molecules," The Chemical Catalogue Co., New York, 1929.
19. C. P. Poole and H. A. Farach, Relaxation in Magnetic Resonance, Academic Press, 1971.
20. F. Perrin, "Etude Mathematique du Mouvement Brownien de Rotation," Ann. Ecole Norm. Sup., Vol. 45, (1928), pp. 1-51.
21. R. F. Fox, Physical Applications of Multiplicative Stochastic Processes, J. of Mathematical Physics, Vol. 14, No. 1, Jan. 1973, pp. 20-25.
22. V. A. Tutubalin, "Multimode Waveguides and Probability Distributions on a Symplectic Group," Theory of Probability and Its Applications, Vol. 16, No. 4, pp. 631-642, 1971.
23. G. F. Carrier, "Stochastically Driven Dynamical Systems," J. of Fluid Mechanics, Vol. 44, Part 2 (1970) pp. 249-264.
24. Martin Clark, (This Volume).
25. R. W. Brockett and J. C. Willems, "Average Value Criteria for Stochastic Stability," Stability of Stochastic Dynamical Systems, (ed. Ruth Curtain), Springer-Verlag, 1972.
26. H. P. McKean, Stochastic Integrals, Academic Press, New York, 1969.
27. S. Helgason, Differential Geometry and Symmetric Spaces, Academic Press, New York, 1960.
28. M. E. Rose, Elementary Theory of Angular Momentum, J. Wiley, New York, 1957.
29. L. Hormander, "Hypoelliptic Second Order Differential Equations," Acta. Math., 119: 147-171, 1967.
30. M. G. Krein, "Introduction to the Geometry of Indefinite J-Spaces and to the Theory of Operators in Those Spaces," Am. Math. Soc. Translations, (Series 2) Vol. 93, 1970, pp. 103-176.
31. I. M. Gel'fand and V. B. Lidskii, "On the Structure of the Regions of Stability of Linear Canonical Systems of Differential Equations with Periodic Coefficients," Am. Math. Soc. Transl. Series 2, Vol. 8, (1958) pp. 143-182.
32. R. W. Brockett and A. Rahimi, "Lie Algebras and Linear Differential Equations," in Ordinary Differential Equations, (L. Weiss ed.) Academic Press, New York, 1972, pp. 379.

1. INTRODUCTION

Electrical networks utilizing electronic switching are used to obtain a variety of results which are difficult or impossible to obtain with conventional linear circuits. Typical applications include circuits which perform elementary control functions, circuits for DC to DC voltage conversion, circuits for frequency conversion, etc. We will be mostly concerned with networks designed for their power handling capability. A reasonably complete

---

*Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts 02138.

**Central Research Laboratories, Texas Instruments Incorporated, Dallas, Texas.

†This research was supported in part by the U.S. Office of Naval Research under the Joint Services Electronics Program by Contract N00014-67-A-0298-0006 and by the National Aeronautics and Space Administration under Grant NGR 22-007-172.

DATA MISSING

storage elements play a significant role. The main thrust of our paper is in this direction.

The most natural description of the basic equations of motion for circuits of interest here is a set of first order differential equations describing the time evolution of the inductor currents and the capacitor voltages. We find that in many cases networks containing diodes, controlled switches, linear time-invariant inductors, capacitors and resistors, current and voltage sources, can be modeled by equations of the form

$$\dot{x}(t) = A_o + \sum_{i=1}^{m} u_i(t)A_i)x(t) + b_o + \sum_{i=1}^{m} u_i(t)b_i$$

where $u_i(t)$ model the effects of switches. Because the right side of this equation contains a set of bilinear terms, it is often referred to as a

bilinear system. It serves as the starting point for the analysis of the dynamical behavior of the electrical network. Next we introduce a set of approximations -- based on averaging -- which yield a family of bilinear equations in which the average values of the switch variables occur. The approximate systems can then be used to design control strategies, either based on linearlization and conventional frequency response or else some of the new stabilization methods introduced here in section 5. The key step is the basic averaging approximation and the various refinements of it. It is at this point and in the treatment of bilinear equations that we make some use of Lie theory.

The paper concludes with the detailed analysis of an example illustrating the use of the techniques discussed in the design of pulse-width modulated regulators for DC to DC conversion systems.

2. MODELS FOR DIODES AND CONTROLLED SWITCHES

In the later sections we will want to model all circuits under consideration by circuits which contain only linear time invariant inductors, capacitors, and resistors, sources and ideal switches. By an ideal switch we understand a circuit element which either transmits no current, regardless of the voltage drop across it, or else has no voltage drop across it regardless of the current through it and it can change from one of these modes to the other on command -- regardless of the current it is carrying.

Our first point is that ideal switches are actually good models for a large number of circuits which are easily built. It is true that silicon controlled rectifiers and power transistors are only approximated by ideal switches in certain regimes because of their turn-on and turn-off dynamics, the fact that they conduct in one direction only, etc. However, there are simple circuits for making these devices bidirectional such as the one shown in figure 1. This figure shows a diode bridge with a unidirectional device in the middle. The overall circuit will conduct current in both directions and can therefore be modeled by a simple switch. The

Figure 1: A bidirectional switch made with 4 diodes and a unidirectional switch, together with its equivalent.

turn-on and turn-off dynamics, while important in some applications, will be ignored here. We justify this on the grounds that these effects are second order compared with the analysis considered here. The restriction to

bidirectional devices is justified on the grounds that the presence of unidirectional devices would complicate our analysis and furthermore they can be eliminated -- sometimes with good effect on system performance -- via the circuit in figure 1 or some modification of it.

We also want to consider three terminal switches of the form represented by figure 2. These can be realized in various ways depending on whether or not fully bidirectional behavior is required. A more or less typical example is the circuit shown in figure 3. It is accurately modeled by

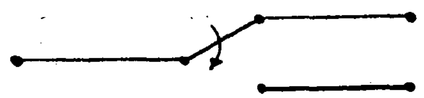

Figure 2: A three terminal switch.

replacing the controlled switch and the diode by a three terminal switch provided the current is always flowing through the inductor in the positive sense.

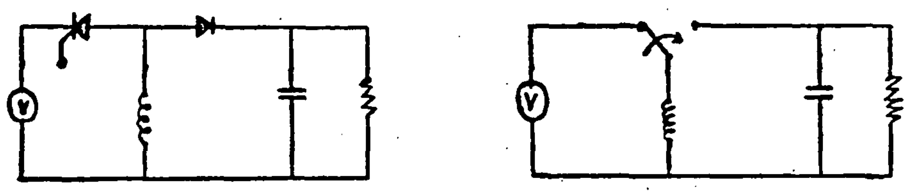

Figure 3: A nonlinear network and a controlled switch model incorporating a three terminal switch.

A general approach to modeling circuits with switches cannot be based on impedance descriptions since the ideal switches do not have impedance

characterizations. However a scattering variable description of a switch does exist and it is desirable to base the whole approach to modeling networks with switches on a scattering variable formulation. That is, if we have a network with switches and sources we extract the switches and relate i+v and i-v across the switch via

$$(i+v) = u(t)(i-v)$$

When the switch is closed i+v = i-v and when it is open i+v = -(i-v). Thus the switch value, u(t), is either plus or minus one.

Based on the above discussion we claim that a good understanding of networks of the form shown in figure 4 would contribute to our ability to design useful circuits. Moreover one easily sees that subject to the basic

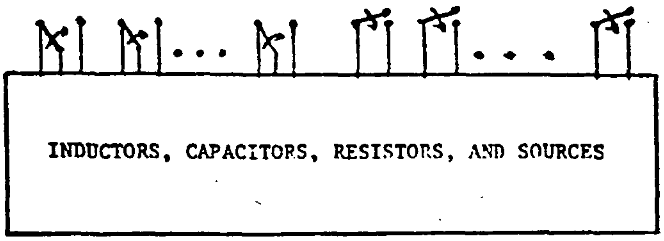

Figure 4: The general time invariant switchable electrical network with two and three terminal switches.

well-posedness conditions, such as arise out of the necessity of not having any capacitive loops or inductor cut sets regardless of the switch configuration, such circuits always yield state equations of the form

$$\dot{x}(t) = (A_o + \sum_{i=1}^{m} u_i(t)A_i)x(t) + b_o(t) + \sum_{i=1}^{m} u_i(t)b_i(t)$$

We refer the reader to standard sources for the justification of this remark. See [1] for background and references.

3. CYCLIC PROCESSES AND VECTOR FIELDS

In section 4 we will discuss certain simplifications for commutated electrical networks which operate in a quasi-periodic mode. In order to motivate the type of analysis which is carried out there we want to discuss the role of switches from a particular point of view which has to do with the cyclic nature of the processes in question.

The idea is illustrated with the circuit shown in figure 3. This circuit can be regarded as a model for a simple voltage converter. We consider the time evolution of the inductor current and capacitor voltage. In terms of these coordinates we have the two different types of integral curves, depending on the switch position. The choices, are shown in figure 5. By following alternatively the paths we can effectively stand

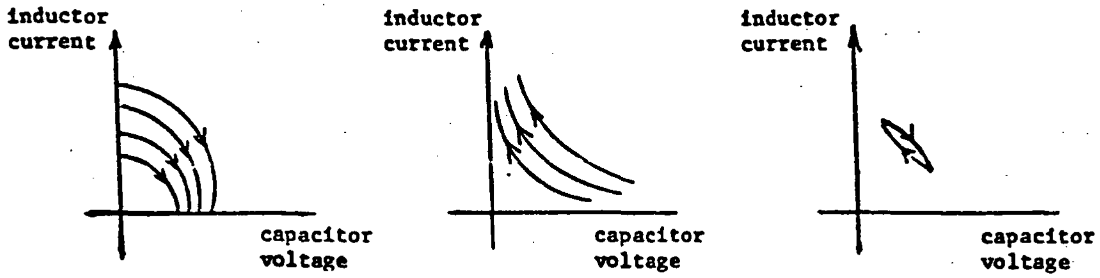

Figure 5: (a) switch closed right; (b) switch closed left; (c) typical cycle

still at an average position. There are, at the next level of complication, circuit effects which cannot be explained on the basis of simple averaging.

In general what one does is to allow the system to follow alternate paths (vector fields) in a definite order, thereby creating effects which are not achievable by following any one of the fixed available paths. This point of view is the basis for studying the controllability of nonlinear systems. See the recent paper of R. Hirschorn [7] for a systematic account.

4. THE MATHEMATICAL ANALYSIS

We now turn to the analysis of circuits modeled by equation (1). This will require certain results from the theory of linear systems and Lie algebras. All necessary background material can be found in [2] and [3].

If A and B are n by n matrices then we use the bracket symbol [A,B] to denote the commutator product

$$[A,B] = AB-BA$$

a subset of the space of n by n matrices which is, (a) a linear space, and (b) contains [A,B] whenever it contains A and B, is called a matrix Lie algebra. If A and B are any two n by n matrices then we can find the smallest Lie algebra which contains them simply by forming their commutator product, taking linear combinations, more commutator products, etc. This process stops in a finite number of steps because the set of n by n matrices is finite dimensional. We denote the smallest Lie algebra which contains A, B, ..., C by $\{A,B,\ldots,C\}_{LA}$.

Given a differential equation system

$$\dot{x}(t) = (A+u(t)B)x(t) + u(t)b + a \tag{1}$$

it is known (see, e.g. [3]) that the transition matrix $\Phi_A(t)$ must belong to the set

$$\{\exp\{A,B\}_{LA}\}_G = \{X\colon X = e^{L_1}e^{L_2}e\ldots e^{L_m},\quad L_i \in \{A,B\}_{LA}\}$$

Moreover, it is known that "fairly large" subsets of $\{\exp\{A,B\}_{LA}\}_G$ can in fact be achieved. See the recent paper of R. Hirschorn [7] and his references to the work of Jurdjevic and Sussmann.

Suppose that we have a linear time invariant system of the form

$$\dot{x}(t) = A_o x(t)+B_o u(t); \quad y(t) = C_o(t)$$

with (A_o,B_o,C_o) a controllable and observable (minimal) system. If $A(\cdot)$, $B(\cdot)$ and $C(\cdot)$ are periodic functions of time such that $||A(\cdot)-A_o||$, $||B(\cdot)-B_o||$ and $||C(\cdot)-C_o||$ are all less than ε. Then for ε sufficiently small the periodic system

$$\dot{x}(t) = A(t)x(t)+B(t)u(t); \quad y(t) = Cx(t)$$

will be minimal as well. The input-output map for the periodic system, i.e. the map $u \longmapsto y$ defined by

$$y(t) = \int_0^t C(t)\Phi_A(t,\sigma)B(\sigma)d\sigma$$

will be close* to that of the time invariant system provided that both are asymptotically stable. This suggests that one might replace the periodic system by a time invariant one obtained by averaging over one period provided the resulting system is both stable and minimal. There are two basic properties of any such approximation which one should demand: (a) if we replace t by $t+\alpha$ the approximation should not change, and (b) A_o should belong to the Lie algebras generated by $\{A(t_i)\}$, so that the "average" system is not exhibiting behavior that the original system could not duplicate for any choice of u. This is significant, for example, if one is to avoid pitfalls such as approximating a lossless network with a lossy one, etc.

---

*Say with respect to the operator topology induced by letting u and y belong to $L_2[0,\infty)$.

The approximation based on simple averaging becomes less satisfactory as the variation about the average becomes larger or as the period of the variation becomes larger. In these cases one wants to refine this approximation further. We describe how this can be done in the piecewise constant case. Notice that given two real matrices A and B there is no guarantee that there exists a real matrix C such that

$$e^A e^B = e^C$$

This will be the case however, if A and B are small enough and then C will be given by a convergent infinite series

$$C = A+B+[A,B] + \frac{1}{12}[[A,B],B] + \frac{1}{12}[[B,A],A] + \ldots$$

known as the Baker-Campbell-Hausdorff formula. Various expressions like this actually form the basis for a large number of useful approximations in physics [4]. What we find here is that they can be quite useful in the analysis of electrical networks as well.

The idea is this. Suppose that A(t) is given by A(t) = A(t+T) and

$$A(t) = \begin{cases} A & 0 \leqslant t \leqslant \alpha \\ B & \alpha \leqslant t \leqslant T \end{cases}$$

Then we want an approximation for $e^{A\alpha}e^{B(T-\alpha)}$. The Baker-Campbell-Hausdorff formula gives such a result, namely e^C where C is as above. Now if we want a formula for C which is independent of order, i.e. insensitive to a shift of origin of the time axis, then we must drop out the [A,B] term to get

$$\ln \frac{1}{2}(e^A e^B + e^B e^A) \simeq A+B + \frac{1}{12}[[A,B]B] + \frac{1}{12}[[B,A],A]$$
$$= A+B + \frac{1}{12}[[A,B],B-A]$$

Putting these ideas together we obtain a series of approximations for piecewise constant periodic systems. If the system equations are

$$\dot{x}(t) = A(t)x(t)+Bu; \quad y = Hx$$

with $A(t+T) = A(t)$ and $A(t)$ as above, then the first approximation is

$$\dot{x}(t) = [\frac{\alpha}{T} A + (1-\frac{\alpha}{T})B]x(t)+Gu(t); \qquad y(t) = Hx(t)$$

The second approximation is

$$\dot{x}(t) = [\frac{\alpha}{T} A+(1-\frac{\alpha}{T})B+\frac{1}{12}(\frac{\alpha}{T} - \frac{\alpha^2}{T^2})[[A,B],(1-\frac{\alpha}{T})B-\frac{\alpha}{T}A]]x(t)+Gu(t);$$

$$y(t) = Hx(t)$$

and higher degrees of approximation can be generated by taking more terms in the Baker-Campbell-Hausdorff formula.

Notice that the inhomogeneous equation

$$\dot{x}(t) = A(t)x(t)+b(t)$$

can be converted to the homogeneous form

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ 1 \end{bmatrix} = \begin{bmatrix} A(t) & b(t) \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x(t) \\ 1 \end{bmatrix}$$

by the simple device of augmenting the x-vector. We will use this trick when it is necessary to approximate the solutions of

$$\dot{x}(t) = Ax(t)+u(t)Bx(t)+u(t)b$$

in subsequent sections, since it allows us to use the Baker-Campbell-Hausdorff formula directly.

5. PULSE-WIDTH MODULATED SYSTEMS

By a pulse-width modulated system we understand a special type of bilinear system of the form

$$\dot{x}(t) = (A+\sum_{i=1}^{m} u_i(t)B)x(t)+\sum_{i=1}^{m} u_i(t)b_i+c$$

where each $u_i(t)$ can take on only two values. Moreover, there is a basic pulse period for the system and $u_i(t)$ can only switch between its two possible values one time in each period. Thus if u switches between one and zero then a typical u(·) looks as shown in figure 6.

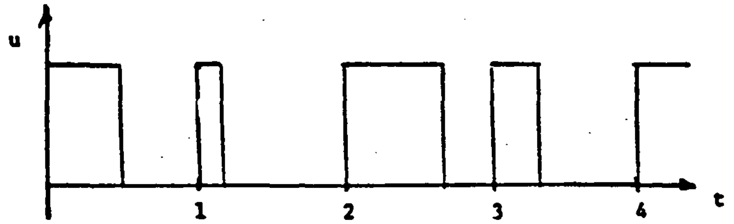

Figure 6: A typical u(·) for a pulse-width modulated system of pulse period equal to one.

A pulse width modulated system is a bilinear system and if one averages over one period the averaged pulse-width modulated system is also a bilinear system. The significant difference is that whereas the original system has controls which take on only two values the averaged system has controls which take on a continuum of values. This change makes the control problem very much easier to study using conventional techniques. It also allows one to apply the theory of bilinear systems [5].

6. STABILIZATION

In section 7 we want to show, by example, that the methods of this paper can be useful in understanding pulse width modulated control systems. However, we also want to indicat how these methods can be used in design. For that reason we examine the question of stabilizing bilinear systems by feedback.

Consider the bilinear system

$$\dot{x}(t) = Ax(t)+u(t)Bx(t)+bu(t)$$

There are 3 more or less obvious remarks to be made about the existence of feedback control laws which make $x = 0$ asymptotically stable.

(i) If $b, Ab, \ldots A^{n-1}b$ is of full rank then by virtue of the pole relocation theorem for linear systems there exists a linear feedback control $u = -k'x$ such that $A-bk$ has all its eigenvalues in $\operatorname{Re} s < 0$. Since $k'xBx$ is quadratic in x it follows that $x=0$ is (locally) asymptotically stable for

$$\dot{x}(t) = Ax(t)-k'x(t)Bx(t)-bk'x(t)$$

(ii) If $\dot{x}(t) = Ax(t)$ is asymptotically stable then the control law $u \equiv 0$ results in stability. If $\dot{x}(t) = Ax(t)$ is stable but not asymptotically stable then there exists a nonsingular symmetric matrix Q such that $QA+A'Q = 0$ and the control

$$u(t) = -x'(QB+B'Q)x-cx(t)$$

gives asymptotic stability unless u vanishes identically for some non-decaying solution of $\dot{x}(t) = Ax(t)$.

(iii) If there exists a choice of basis such that

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}; \quad B = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}; \quad C = \begin{bmatrix} C \\ 0 \end{bmatrix}$$

and if

$$\dot{x}(t) = A_{22}x(t)+u(t)B_{22}x(t)$$

meets the conditions for instability for the circle theorem [2] or some other instability criterion then there is no stabilizing control law.

We now describe a refinement of (ii) which is well suited for the study of the electrical network problems we have been discussing.

Theorem: Suppose that A is similar to a skew symmetric matrix and suppose that the eigenvalues of A satisfy the condition $\lambda_i+\lambda_j \neq \lambda_k$ for all i, j and k. Assume that b, Ab,...,$A^{n-1}b$ are linearly independent. Then there exists an n by n matrix Q with Z = Q' > 0 such that QA+A'Q = 0 and the control law

$$u = -x'(QB+B'Q)x-x'Qb$$

makes the null solution of

$$\dot{x}(t) = Ax(t)+u(t)Bx(t)+bu(t); \quad x(t) \in \mathbb{R}^n$$

asymptotically stable in the large.

Proof: The existance of a Q satisfying the given condition is classical (see e.g. [2]). We want to use the Liapunov function x'Qx and a theorem of LaSalle which gives asymptotic stability if x'Qx is positive definite and $\dot{v}$ is not identically zero along a nonzero trajectory. In this case

$$\dot{v} = -u^2$$

Now if $\dot{v}$ is identically zero then u is zero and $\dot{x}$ = Ax. We also have

$$x'QB+QB'x = -x'Qb$$

but if u vanishes then $x(t) = e^{At}x_o$. Since $QA = -A'Q$ we see $x'e^{A't}Qb = x'Qe^{-At}$ and, by the condition on $b, Ab, \ldots, A^{n-1}b$ this does not vanish identically for $x_o \neq 0$. Thus we must have

$$x_o'e^{A't}(QB+B'Q)e^{At}x_o = -x_o'e^{A't}Qb$$

with both sides nonzero. Now the left side is a sum of terms of the type $\sigma_i e^{\lambda_i t}$ and confluent forms where λ_i are the eigenvalues of A. The right side is a similar sum of $\beta_i e^{(\lambda_i+\lambda_j)t}$. By hypothesis the exponentials on the right and left are distinct. Since the left side does not vanish identically they are not equal.

It may be that the conclusion holds under the weaker condition that there should exist no vector $x \neq 0$ such that $ax+x'(QB+B'Q)x(Bx+b) = 0$. This condition is certainly necessary and perhaps it is sufficient as well.

7. EXAMPLES

This section consists of an example illustrating the application of the analysis done in sections 4 and 6. We are especially interested in determining the effects of going to higher order approximations.

We consider the network shown in figure 3b. The equations of motion are, assuming a one volt supply with positive polarity down we have

$$\begin{bmatrix} L\dot{x} \\ C\dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1-u \\ u-1 & -R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad (*)$$

Here x is the inductor current, v is the capacitor voltage, and u=1 when the switch is closed on the left and 0 when it is closed on the right. If we assume u is operated periodically then the first approximation is

$$\begin{bmatrix} L\dot{z}_1 \\ C\dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1-\bar{u} \\ \bar{u}-1 & -R \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \bar{u} \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad (**)$$

where $\bar{u}$ is the average value of u. We can associate a time invariant network with these equations in various ways. If we want to preserve the meaning of z_2 as the voltage across the resistor then the network in figure 7 is appropriate.

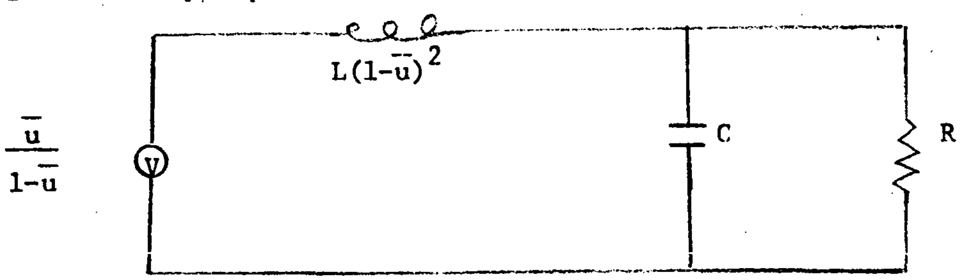

Figure 7: Network equivalent of the averaged equation.

If u is periodic and of the form shown in figure 8 then we can refine this approximation by taking additional terms in the Baker-Campbell-Hausdorff formula. Introduce A and B, taken from equation (*), as

$$A = \begin{bmatrix} 0 & 1 \\ -1 & -R \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

The next approximation is then

$$\ell n \frac{1}{2} \{ e^{\alpha(A+B)} e^{(1-\alpha)A} + e^{(1-\alpha)A} e^{\alpha(A+B)} \}$$

$$\simeq A + \alpha B + \frac{1}{12} \{ \alpha(1-\alpha)(1-2\alpha)[[B,A],A] + \alpha^2(1-\alpha)[[A,B],B] \}$$

$$\simeq A + \alpha B + \frac{1}{12} \ \alpha(1-\alpha)(1-2\alpha) \begin{bmatrix} -2R & R^2 \\ -R^2 & +2R \end{bmatrix} + \alpha^2(1-\alpha) \begin{bmatrix} 2R & 0 \\ 0 & -2R \end{bmatrix}$$

Near a 50% duty cycle the second correction term is more important than the first, which, in fact, vanishes at $\alpha = 1/2$. The second term has the effect

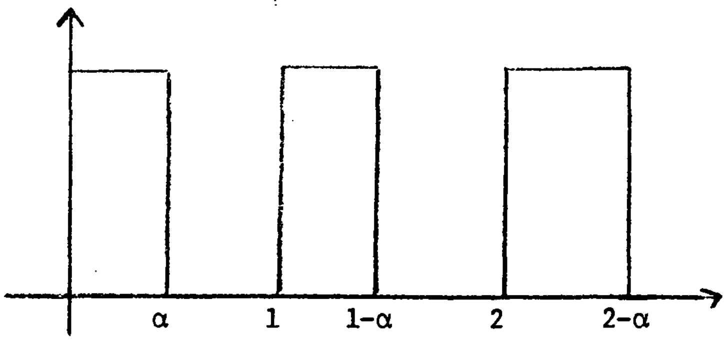

Figure 8: The duty cycle for switch in figure 3.

of decreasing the output resistor by $\alpha^2(1-\alpha)$ R/12 and inserting this same value of resistance in series with the inductor as shown in figure 9. The actual percentage change in the output resistor is about 2% but this together with the insertion of the small resistor in series with the inductor has a notable effect on the frequency response characteristics.

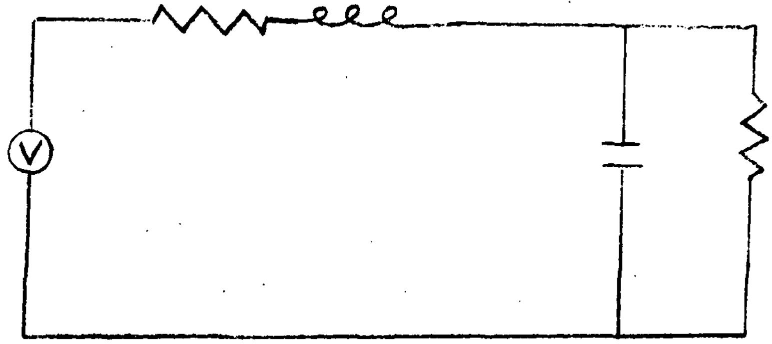

Figure 9: A better approximation for the network in figure 3.

Based on the first approximation described by (**) let's look for a control law which stabilizes the output voltage to a value $z_2^* > 0$. This means that for steady state we must have

$$\begin{bmatrix} 0 & 1-\bar{u}_o \\ \bar{u}_o-1 & -R \end{bmatrix} \begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} + \bar{u}_o \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This fixed the values of z_1^* and $\bar{u}_o$

$$\bar{u}_o = \frac{z_2^*}{z_2^*+1}$$

$$z_1^* = \frac{Rz_2^*}{\bar{u}_o-1} = -Rz_2^*(z_2^*+1)$$

Say that $z_2^* = 4$ and $R = 1$. Then $u_o = 4/5$ and we want to stabilize at

$$\begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix} = \begin{bmatrix} -20 \\ 4 \end{bmatrix}$$

Introduce $z-z^* = y$ and $\bar{u}-\bar{u}_o = v$. In these coordinates we have

$$\begin{bmatrix} L\dot{y}_1 \\ C\dot{y}_2 \end{bmatrix} = \begin{bmatrix} 0 & -1/5 \\ -1/5 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + v \begin{bmatrix} 0 & -1 \\ +1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + v \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

If we now linearize this equation about $y = 0$, $v = 0$ we see that the transfer function between the output voltage error

$$y_2 = (z_2-4)$$

and the deviation of the input average value from 4/5

$$v = \bar{u} - 4/5$$

is

$$\hat{y}_z(s)/\hat{v}(s) = \frac{1/5L}{Cs^2+s+1/25L}$$

This linearization is obtained by ignoring the vy_1 and vy_2 terms. This is the relevant transfer function for feedback regulator design, regarding the average switch position as the input and the deviation of the output voltage from its average value as the output. If necessary one can now return to the refined approximation and work out a more accurate equivalent model.

8. CONCLUSIONS

We have shown here that commutated electrical networks can be analyzed in an approximate way by using an averaging technique based on first order differential equation descriptions and the Baker-Campbell-Hausdorff formula from Lie theory. There are three steps in the analysis:

(a) replace all unidirectional switches by bidirectional equivalents or bidirectional approximations valid in the operating regime,

(b) introduce equivalent circuit equations based on averaging and expansion of $e^{A}e^{B}$,

(c) stabilize the resulting bilinear equations using linearization or bilinear theory.

An example is given to indicate the type of insight available from this approach.

REFERENCES

1. C.A. Desoer and E.S. Khu, Basic Circuit Theory, McGraw-Hill, N.Y., 1969.

2. R.W. Brockett, Finite Dimensional Linear Systems, J. Wiley, N.Y., 1970.

3. R.W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM J. on Control, Vol. 10, 1973, pp. 265-284.

4. R.M. Wilcox, "Exponential Operators and Parameter Differentiation in Quantum Physics," J. of Math. Physics, Vol. 9, 1967, pp. 962-982.

5. R.W. Brockett, "On the Algebraic Structure of Bilinear Systems," in Theory and Application of Variable Structure Systems, (R.R. Mohler and A. Ruberti, eds.) Academic Press, N.Y., 1972.

6. J. Wood, Power Conversion in Electrical Networks, Ph.D. Thesis, M.I.T., Dept. of Electrical Engineering, 1973.

7. R. Hirschorn, "Control of Bilinear Systems," in IEEE Special Publication, Monographs in Electrical Network Theory, (C.A. Desoer ed.), IEEE, New York, 1974.

IV. SUMMARY AND CONCLUSIONS

R.W. Brockett

1. Main Results

NASA support has resulted in the work leading to the publications [1-25] cited in the reference list. While a detailed summary of this work is obviously impossible there are several main lines of thought which are apparent. We list these with an indication of their origin.

1. High efficiency power conversion networks are usually well approximated by electrical networks with linear inductors, capacitors, resistors (for load), sources, ideal switches and diodes, together with the control circuitry for the switches. [1-5]

2. Switched electrical networks of this type are, without the control circuitry, bilinear systems. [1,3,4]

3. Many aspects of the behavior of bilinear systems can be understood on the basis of mathematical models without resorting to simulation. The study of controlled bilinear systems such as arise with regulated DC to DC supplies in which linear or nonlinear feedback is applied to a bilinear system is more difficult but design for stability is possible based on mathematical models. [2,5]

4. The use of Lie algebraic techniques is essential for understanding the controllability and observability of bilinear systems. Moreover, these same techniques carry over with little change to more general nonlinear systems. This last point is important in understanding the feedback control of bilinear systems since even linear feedback leads to systems which are no longer bilinear. [2,3]

5. The standard use of averaging to approximate the behavior of switching regulators can be refined using Lie algebraic techniques. This refinement is useful when the natural frequencies of the regulator and the clock frequency are not widely separated. [5]

Taken together the methods developed here constitute a basis for understanding some of the theoretical problems which arise in the study of power conversion networks. The idea that Lie algebras shed some light on the control of switchable electrical networks is felt to be one of the major contributions. A priori there was no hint in the literature that this might be true. Equally important is the idea that the basic tools of modern control theory, e.g. state space models, Liapunov stability methods, optimal control, etc., can be of practical value in designing control laws for converters and regulators. To be sure, this latter idea is becoming widely recognized, and our work only reinforces an established trend. However our mathematical methods can only go so far toward the solution of the design problems and further interactions with system designers should be useful in refining the methodology generated so far.

2. Difficulties

The principle remaining difficulties in analyzing nonlinear equations of the type which occur in power processing problems lie in the area of:

1. Free running converters for which the clock speed is not a priori fixed but which depends on the load and supply conditions.

2. Converters which face widely changing load conditions such as would cause the system to change its basic mode of operations.

3. The design of multimode converters controlled by finite state systems of considerable complexity.

For those applications in which reliability considerations outweigh cost, it seems likely that more sophisticated digital control circuitry will become more common. This will make item three very important in this context. It also seems likely that an increasing number of applications will be found where efficiency is the overriding consideration and for these cases sophisticated digital control circuitry may be justified also.

Though it has been recognized for a long time that there is a real need for a theory of systems which are partly continuous and partly finite state, results have been slow in coming. It may be that previous efforts have addressed the problem in too much generality and have not exploited the special features of the known successful applications. In any case this problem seems too important to ignore in spite of the apparent difficulty.

3. Future Work

The main hope for further simplification in this area of nonlinear analysis rests in finding a suitable extension of the transfer function idea. Recent work on Volterra series* indicates that the Volterra kernels for the input-output map for systems governed by ordinary differential equations can be computed rather easily and can be of use in understanding the behavior of systems. This idea has already been worked out in detail by d'Alessandro, Isidori, and Ruberti** for bilinear systems and it seems to hold great promise for future developments.

It is also clear that more work should be done which recognizes explicitly the role of logic elements in the controller. This is a difficult problem area but one of great importance.

Finally, in view of the great importance which one must place on efficiency it seems that more emphasis should be placed on the development of fundamental bounds on efficiency. We feel that the work of Wolaver† on fundamental limitations on converter circuits is an excellent start and that this line of work deserves more attention.

---

*R.W. Brockett, "Volterra Series and Geometric Control Theory," Proc. International Federation on Automatic Control, Boston, Mass. 1975.

**P. d'Alessandro, A. Isidori, and A. Ruberti, "Realizations and Structure Theory of Bilinear Dynamical Systems," SIAM J. on Control, (to appear).

†D. Wolaver, Fundamental Study of DC to DC Conversion, Ph.D. Thesis, M.I.T., Dept. of Electrical Engineering, June 1969.

4. References

Reports

1. J.R. Wood, "Power Conversion in Electrical Network," NASA Report CR-120830.

2. R.W. Brockett et al., "Algebraic Methods in System Theory," NASA Report CR-121170.

3. R.W. Brockett et al., "Differential Geometric Methods in Control," NASA Report CR-120946.

4. R.W. Brockett, "Network Synthesis," NASA Report CR-121171.

5. R.W. Brockett and J. Wood, Final Report, "Application of System Theory to Power Processing Problems," NASA Report CR-134708. (this report)

The following papers appearing in the technical literature acknowledge the support of this grant. In some cases they form the basis for parts of the above reports.

6. J.L. Willems, "Space Harmonics in Unified Electrical-Machine Theory," Proc. IEE, Vol. 118, No. 10, pp. 1408-1412, October 1971.

7. R.W. Brockett and A. Rahimi, "Lie Algebras and Linear Differential Equations," Ordinary Differential Equations, L. Weiss editor, Academic Press, 1972, pp. 379-386.

8. R.W. Brockett, "Decomposition Methods for Nonlinear Systems," in IEEE Special Publication, System Structure, 1971 (S. Morse ed.).

9. R.W. Brockett, "Differential Geometric Methods in System Theory," Proc. IEEE Decision and Control Conference, Dec. 1971, pp. 176-180.

10. R.W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM J. Control, Vol. 10, No. 2, May 1972, pp. 265-284.

11. R.W. Brockett and J.C. Willems, "Average Value Criteria for Stochastic Stability," Stability of Stochastic Dynamical Systems, (ed. Ruth Curtain), Springer-Verlag Lecture Notes on Mathematics, Vol. 294, 1972.

12. R.W. Brockett and A. Rahimi, "Homotopic Classification of Minimal Periodic Realizations of Stationary Weighting Patterns," SIAM J. Appl. Math., Vol. 22, No. 3, May 1972, pp. 481-486.

13. R.W. Brockett and A.S. Willsky, "Finite State Group Homomorphic Sequential Machines," IEEE Trans. on Automatic Control, Vol. AC 17, 1972, pp. 483-490.

14. R.W. Brockett, "On the Algebraic Structure of Bilinear Systems," in Variable Structure Systems, Academic Press, N.Y., 1972, pp. 153-168, (R. Mohler and A. Ruberti eds.).

15. V. Jurdjevic and H.J. Sussmann, "Control Systems on Lie Groups," Journal of Differential Equations, Vol. 12, No. 2, Sept. 1972, pp. 313-329.

16. H.J. Sussmann and V. Jurdjevic, "Controllability of Nonlinear Systems," J. of Differential Equations, Vol. 12, No. 1, July 1972, pp. 95-116.

17. R.W. Brockett, "Lie Theory and Control Systems Defined on Spheres," SIAM J. on Applied Math., Vol. 25, No. 2, Sept. 1973, pp. 213-225.

18. R.W. Brockett, "Lie Algebras and Lie Groups in Control Theory," Geometric Methods in System Theory, Reidel Publishing Co., (D.Q. Mayne and R.W. Brockett eds.), 1973.

19. R.M. Hirschorn, "Controllability in Nonlinear Systems," Geometric Methods in System Theory, Reidel Publishing Co., (D.Q. Mayne and R.W. Brockett eds.), 1973.

20. R.M. Hirschorn, "Topological Semigroups, Sets of Generators, and Controllability," Duke Mathematical Journal, Vol. 40, No. 4, December 1973, pp. 937-947.

21. D. Williamson, "Oscillations in Nonlinear Feedback Systems," Proc. 1973 Joint Automatic Control Conference, Columbus, Ohio, 1973.

22. R.W. Brockett and J. Wood, "Electrical Networks Containing Controlled Switches," in IEEE Special Publication, Applications of Lie Group Theory to Nonlinear Network Problems, (C.A. Desoer ed.), Western Periodicals Co., North Hollywood, California, 1974.

The following Harvard theses acknowledge partial support of the grant.

23. D. Williamson, A Geometric Theory of Oscillations in Non Linear Feedback Systems, Ph.D. Thesis, Division of Engineering and Applied Physics, Harvard University, Sept. 1973.

24. R.M. Hirschorn, Topological Semigroups and Controllability in Bilinear Systems, Ph.D. Thesis, Division of Engineering and Applied Physics, Harvard University, Sept. 1973.

25. F.N. Parr, Finitely Specified Infinite Dimensional Systems, Ph.D. Thesis, Division of Engineering and Applied Physics, Harvard University, September 1974.

REPORT DISTRIBUTION LIST

National Aeronautics & Space Administration
Lewis Research Center
21000 Brookpark Road
Cleveland, Ohio 44135

			COPIES
Attn:	H.W. Plohr	M.S. 54-1	1
	P.A. Thollot	M.S. 54-1	1
	R.R. Secunde	M.S. 54-1	1
	R.C. Finke	M.S. 54-1	1
	V.R. Lalli	M.S. 500-211	3
	J.E. Triner	M.S. 54-1	1
	N.T. Musial	M.S. 500-311	1
	G. Mandel	M.S. 6-2	1
	Librarian	M.S. 60-3	2
	Report Control Office	M.S. 5-5	1
	J.W. Miser	M.S. 3-10	1
	Technology Utilization Office	M.S. 3-19	1
	J.E. Dilley	M.S. 500-313	1
	Reliability & Quality Assurance	M.S. 500-211	1
	C.S. Corcoran, Jr.	M.S. 500-201	

National Aeronautics & Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

Attn:	Dr. D. Harris	M.S. 11-E17	1
	D.R. Pasciutti	M.S. 11-E43	1
	C.M. MacKenzie	M.S. 11-C1	1
	A. Lunchick	M.S. 6-S41D	1
	W. Isley	M.S. 22-254	1
	Librarian		1

National Aeronautics & Space Administration
Lyndon B. Johnson Space Flight Center
Houston, Texas 77058

Attn:	A.B. Eickmeier (EB)		1
	F.E. Eastman (EB3)		1
	W.R. Dusenberry (EP5)		1
	Librarian		1

			COPIES
National Aeronautics & Space Administration Marshall Space Flight Center Huntsville, Alabama 35812			
Attn:	Gene Young, S&L-ASTR-EPN		1
	Roy Lanier, S&E-ASTR-EPN		1
	D.W. Westrope, PD-PP-E		1
	C.H. Guttman, PD-SA-P		1
	J.P. Hethcoat, PD-SA-P		1
	Librarian		1
	J. Felech, S&E-ASTR-EB		
National Aeronautics & Space Administration Ames Research Center Moffett Field, California 94035			
Attn:	Dr. Hans Mark	M.S. 200-1	1
	Librarian		1
National Aeronautics & Space Administration Langley Research Center Hampton, Virginia 23365			
Attn:	J.L. Paterson	M.S. 477	1
	Merton Sussman	M.S. 488	1
	Robert Wells	M.S. 326	1
	Eugene L. Kelsey	M.S. 326	1
	H.H. Youngblood	M.S. 488	1
	Librarian		1
National Aeronautics & Space Administration Flight Research Center P.O. Box 273 Edwards, California 93523			
Attn:	Clark Hawk		1
	Librarian		1
National Aeronautics & Space Administration John F. Kennedy Space Center Kennedy Space Center, Florida 32899			
Attn:	IS-DOC-IL/Librarian		1

			COPIES
National Aeronautics & Space Administration Washington, D.C. 20546			
Attn:	RP/W.H. Woodward		1
	RD-P/G.W. Cherry		1
	RT/W.S. Aiken, Jr.		1
	RL/H.W. Johnson		1
	BF-RM/F.J. Sullivan		1
	Librarian		1
	RPD/P.R. Miller		1
	RPM/J. Lazar		1
Jet Propulsion Laboratory 4800 Oak Grove Drive Pasadena, California 91103			
Attn:	T.J. Williams	M.S. 198-220	1
	D.J. Kerrisk	M.S. 122-123	1
	E.V. Pawlik	M.S. 122-123	1
	G.W. Wester	M.S. 198-220	1
	Librarian	M.S. 111-113	1
NASA Scientific & Technical Information Facility Attn: Acquisitions Branch P.O. Box 33 College Park, Maryland 20740			30
U.S. Air Force Wright Patterson AFB Dayton, Ohio 45433			
Attn:	Robert Johnson (APIE-2)		1
	George Sherman (AFAPL/POD)		1
	P.E. Stover (AFAPL/POD)		1
U.S. Army Engineer R&D Laboratory Ft. Delvoir, Virginia 22060			
Attn:	R.M. McKechnie, SMEFB-EA		1
U.S. Army R&D Laboratory Ft. Monmouth, New Jersey 07703			
Attn:	H.J. Byrnes, AMSEL-KL-PE		1
	Frank Wrublewski AMSEL-KL-PE		1

			COPIES
Bureau of Naval Weapons Department of the Navy Washington, D.C. 20360			
Attn:	W.T. Beatson (Code RAEE-52)		1
	Milton Knight (Code RAEE-511)		1
Naval Research Laboratory Department of the Navy Washington, D.C. 20390			
Attn:	B.J. Wilson (Code 5230)		1
Naval Air Systems Command Department of the Navy Washington, D.C. 20360			
Attn:	Phil Binderman		1
	AIR 530141C/James C. Taylor		
Power Systems Group Department of the Navy Washington, D.C. Attn: Cmdr. H.L. Lewallen			1
U.S.A. Mibility Equipment Research & Development Center Fort Belvoir, Virginia 22060			
Attn:	E. Reimers		1
Dept. of Transportation Transportation Systems Center 55 Broadway Cambridge, MA 02142			
Attn:	Mr. Frank Raposa	M.S. TMP	1
Advanced Research Project Agency The Pentagon, D.C. 20325			
Attn:	John Huth		1
University City Science Institute Power Information Center Rm. 2107 3401 Market Street Philadelphia, PA 19104			1

	COPIES
Joint DOT-NASA CARD Implementation Office 400 Seventh Street SW Washington, DC 20590 Attn: TST-31/William N. Gardner	1
Aerospace Corporation P.O. Box 95095 Los Angeles, California 90045	
Attn: Library Technical Documents Group	1
Air-Research Manufacturing Company Division of the Garrett Corporation P.O. Box 5217 Phoenix, AZ 85010	
Attn: Mrs. J.F. MacKenzie, Librarian	1
Air-Research Manufacturing Company 2525 W. 190th Street Torrance, California 90509	
Attn: Erwin Detken	1
Librarian	1
Arnold Research Organization, Inc. Von Karman Gas Dynamics Facility Aerodynamics Division Arnold Air Force Sta. TN 37389	
Attn: C.J. Schueler, Chief	1
AUCO Everett Research Lab. 2385 Rever Beach Parkway Everett, MA 02148	
Attn: Dr. J. Zar	1
Bell Aerospace Company Buffalo, N.Y. 14240	
Attn: Alan Coles	1
Bell Telephone Labs. Whippany, New Jersey 07981	
Attn: F.F. Judd	1

	COPIES
General Dynamics Astronautics Dept. 963-2 5001 Kearney Villa Road San Diego, CA Attn: R. Schaelchlin	1
Gen. Dynamics Electro Dynamic Div. P.O. Box 2566 Orlando, FL 32802 Attn: Dr. R.A. Inciardi	1
Gen. Electric Company Missile & Space Div. Valley Forge Space Center P.O. Box 8555 Philadelphia, PA 19101 Attn: J.H. Hayden	1
General Electric Company Solid State & Electronics Lab. Schenectady, N.Y. 12301 Attn: V.L. Stout	1
Gen. Motors Corporation Gen. Motors Defense Research Lab. Santa Barbara, CA Attn: T.M. Corry	1
Honeywell, Inc. Ordance Div. 600 Second Street, N. Hopkins, MI 55343 Attn: Lowell Westbrook	1
Hughes Aircraft Co. Space Systems Div. El Segundo, CA 90245 Attn: D. Garth	1
Hughes Research Laboratories 3011 Mailbu Canyon Rd. Malibu, CA 90565 Attn: W. Herron	1
H. King	1
J. Molitor	1
Lear Siegler, Inc. Power Equipment Division P.O. Box 6719 Cleveland, Ohio 44101 Attn: R.W. Hobel	1

	COPIES
Bendix Corporation Bendix Systems Division Ann Arbor, MI Attn: K.A. More	1
Boeing Company Aerospace System Division P.O. Box 3999 Seattle, WA 98124 Attn: Frank Springgate	1
Bose Corporation 17 Erie Drive East Natick Industrial Park Natick, MA 01760 Attn: Dr. Amar Bose	1
Comsat Laboratories P.O. Box 115 Clarksburg, Maryland 20734 Attn: B. Free	1
Colorado State University Fort Collins, CO 80521 Attn: Dr. P. Wilbur	1
Delco Electronics Gen. Motors Corporation 6767 Hillister Ave. Goleta, CA 93017 Attn: J.F. Britt	1
Duke University School of Engineering Dept. of Electrical Engineering Durham, NC 27706 Attn: Dr. T.G. Wilson	1
Electro-Optical Systems, Inc. 300 N. Halstead Pasadena, CA 91107 Attn: R. Worlock	1
Engineered Magnetics Division Gulton Industries, Inc. 13041 Cerise Ave. Hawthorne, CA Attn: Harry Linden	1
George Lachenmeier	1

	COPIES
Fairchild Industries Fairchild Space & Electronics Div. Germantown, MD 20767 Attn: Mr. D. Robinson	1
Lowell Technological Institute Dept. of Electrical Engineering Lowell, MA 01854 Attn: Dr. Peter Berger	1
Helpar, Inc. 3000 Arlington Boulevard Falls Church, VA Attn: S.S. Wilmarth	1
North American Rockwell Corp. Autonetics Division P.O. Box 4173 3370 E. Anaheim Road Anaheim, CA 92803 Attn: J.W. Adkison	1
Radio Corp. of America ALD Box 800 Princeton, NJ 08540 Attn: P. Pierce	1
TRW, Inc. 23555 Euclin Avenue Cleveland, Ohio Attn: Library	1
TRW, Inc. TRW Systems Group One Space Park Redondo Beach, CA 90278	
Attn: P. Staudhammer	1
E. Cohen	1
A. Schoenfeld	1
J. Biess	1
Westinghouse Electric Corp. Aerospace Electrical Div. Lima, Ohio 45801	1
Westinghouse Electric Corp. Research Labs Churchill Boro Pittsburgh, PA 15235 Attn: Dr. Paul Pittman	1

		COPIES
Highes Aircraft Co. Space and Communications Group El Segundo, CA 90245 Attn: G.I. Cardwell		1
Wilorco, Inc. 729 W. Anaheim Street Long Beach, CA 90813 Attn: Kurt Wilner		1
California Institute of Technology Electrical Engrg. Dept. Room 116-91 Pasadena, CA 91109 Attn: Prof. R.D. Middlebrook		1
Power Electronics Associates 18839 Timberr Lane Fairview Park, Ohio 44125 Attn: Dr. F.C. Schwartz		1
Thornton Power Systems 67 Beaver Street Waltham, MA 02154 Attn: Prof. Dick Thornton D.A. Saar		1 1
Electro-Optical Systems Xerox Corporation Special Products & Systems Div. 300 N. Halstead Street Pasadena, CA 91107 Attn: Dr. S.J. Lindena		1
General Electric Company 1200 Lawrence Parkway Eire, PA 16501 Attn: L.J. Yeager		1
The Boeing Company Commercial Airplane Group P.O. Box 3707 Seattle, WA 98124 Attn: A. Schmidt	M.S. 47-18	1
Martin Marietta Corp. P.O. Box 179 Denver, CO 80201 Attn: R.S. Miller R. Stewart W. Collins		1 1 1 1